МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»

ЛИНЕЙНАЯ АЛГЕБРА

Элементы общей алгебры.
Линейные пространства.
Линейные операторы.
Евклидовы и унитарные пространства.
Билинейные и квадратичные формы.

Учебно-методическое пособие для вузов

Составитель:
К.П. Лазарев

Воронеж 2023

Содержание

I Элементы общей алгебры **6**

1. Группы, кольца, поля **6**

1.1. Группа. Простейшие свойства группы 6

1.2. Подгруппа. Критерий подгруппы. Пересечение подгрупп . 10

1.3. Произведение (сумма) подмножеств группы. Произведение (сумма) подгрупп 11

1.4. Смежные классы группы и их свойства 15

1.5. Конечные группы и их построение. Теорема Лагранжа о порядке подгруппы 17

1.6. Нормальный делитель. Фактор-группа 20

1.7. Кольцо. Простейшие свойства кольца. Подкольцо. Критерий подкольца. 21

1.8. Поле 26

1.9. Кольцо классов вычетов 32

1.10. Изоморфизм групп, колец, полей 35

2. Комплексные числа **37**

2.1. Комплексные числа. Поле комплексных чисел 37

2.2. Алгебраическая форма комплексных чисел. Арифметические операции для комплексных чисел в алгебраической форме. Геометрическое изображение комплексных чисел и операций сложения и вычитания 40

2.3. Тригонометрическая форма комплексного числа. Операции над комплексными числами в тригонометрической форме 42

2.4. Свойства модуля комплексных чисел. Свойства операции сопряжения 48

2.5. Возведение в степень. Извлечение корня 50

3. Многочлены **61**

3.1. Кольцо многочленов от одной переменной 61
3.2. Деление многочленов с остатком. Схема Горнера 67
3.3. Делители многочленов и их свойства 71
3.4. Общий делитель и наибольший общий делитель многочленов. Нахождения наибольшего общего делителя многочленов. Представление наибольшего общего делителя многочленов через эти многочлены 74
3.5. Взаимная простота многочленов. Свойства, связанные со взаимной простотой 81
3.6. Корни многочленов. Кратность корня 85
3.7. Основная теорема алгебры для многочленов над C 88
3.8. Следствия из основной теоремы алгебры для многочленов над полем C 93
3.9. Многочлены над полем R 97

II Линейные пространства. Линейные операторы 101

4. Линейное пространство над произвольным полем 101

4.1. Линейное пространство над произвольным полем, его простейшие свойства. Подпространство. Критерий подпространства 101
4.2. Линейная комбинация векторов. Линейная зависимость и линейная независимость векторов 108
4.3. База и ранг системы векторов. Критерий базы 115
4.4. Базис линейного пространства. Координаты вектора в базисе и их свойства 121
4.5. Размерность линейного пространства. Связь понятий базиса и размерности. Дополнение линейно независимой системы векторов до базиса 125
4.6. Матрица перехода от одного базиса к другому. Связь координат вектора в различных базисах 128
4.7. Линейная оболочка векторов 134
4.8. Сумма и пересечение подпространств 136

4.9. Размерность суммы двух подпространств 140
4.10. Прямая сумма подпространств. Дополнительное подпространство 147
4.11. Линейное многообразие 150
4.12. Изоморфизм линейных пространств 151

5. Линейные операторы 156

5.1. Линейный оператор. Свойства линейного оператора 156
5.2. Матрица линейного оператора. Линейное выражение координат образа вектора при линейном отображении. Связь матриц линейного оператора в различных базисах 159
5.3. Сумма операторов, произведение оператора на число, произведение операторов и соответствующие им матрицы . . . 163
5.4. Образ и ядро линейного оператора 167
5.5. Собственные векторы и собственные значения линейного преобразования 170
5.6. Условия приведения матрицы линейного преобразования к диагональному виду. Жорданова форма матрицы линейного преобразования 178

6. Евклидовы и унитарные пространства 180

6.1. Определение евклидова и унитарного пространства. Простейшие свойства скалярного произведения 180
6.2. Неравенство Коши–Буняковского. Длина вектора 184
6.3. Ортогональность 187
6.4. Метод ортогонализации Грама–Шмидта. Построение ортогонального (ортонормированного) базиса 191
6.5. Ортогональное дополнение. Разложение пространства на сумму ортогональных подпространств. Задача о перпендикуляре 193

7. Билинейные и квадратичные формы 197

7.1. Билинейные формы, их свойства и матрицы 197
7.2. Квадратичные формы, их свойства и матрицы. Канонический вид квадратичной формы. Приведение квадратичной формы к каноническому виду 201

Часть I
Элементы общей алгебры

1. Группы, кольца, поля

1.1. Группа. Простейшие свойства группы

Определение 1. *Множество* $\Gamma \neq \emptyset$ *называется группой, если*

(I) *во множестве* Γ *определена алгебраическая операция* $\circ\colon \Gamma \times \Gamma \to \Gamma$ (*т.е.* $\forall\, u, v \in \Gamma\, \exists!\, w \in \Gamma$, *обозначаемый* $w = u \circ v$);

(II) *выполнены следующие аксиомы*

(АГ1) $\forall\, a, b, c \in \Gamma \quad (a \circ b) \circ c = a \circ (b \circ c)$ — *операция ассоциативна;*

(АГ2) $\exists\, e \in \Gamma\, \forall\, a \in \Gamma \quad a \circ e = e \circ a = a$ — *операция имеет нейтральный элемент* e;

(АГ3) $\forall\, a \in \Gamma\, \exists\, a' \in \Gamma \quad a \circ a' = a' \circ a = e$ — *для каждого элемента* $a \in \Gamma$ *существует симметричный элемент* $a' \in \Gamma$.

(III) *Группа называется коммутативной или абелевой, если выполнена аксиома*

(АГ4) $\forall\, a, b \in \Gamma \quad a \circ b = b \circ a$ — *операция коммутативна.*

Обозначение: группа Γ *или* $\langle \Gamma, \circ \rangle$ — *группа* Γ *с операцией* $\circ$.

Группу Γ *с операцией сложения обозначают* $\langle \Gamma, + \rangle$ *и называют аддитивной, нейтральный элемент называют нулём (обозначение* 0*), симметричный для* a *элемент называют противоположным (обозначение* $(-a)$*).*

Группу Γ *с операцией умножения обозначают* $\langle \Gamma, \cdot \rangle$ *и называют мультипликативной, нейтральный элемент называют единицей (обозначение* e *или* 1*), симметричный для* a *элемент называют обратным (обозначение* a^{-1}*).* ■

Замечание. С помощью метода математической индукции доказывается утверждение: *результат вычисления операции для любого количества элементов группы не зависит от расстановки скобок.* Поэтому скобки можно не ставить.

Примеры: 1) $\langle Z,+\rangle$, $\langle Q,+\rangle$, $\langle R,+\rangle$, $\langle R^{m\times n},+\rangle$ — аддитивные абелевы группы;

2) $\langle Q\setminus 0,\cdot\rangle$, $\langle R\setminus 0,\cdot\rangle$ — мультипликативные абелевы группы;

3) $\langle\{A\in R^{n\times n}|\det(A)\neq 0\},\cdot\rangle$, $\langle\{A\in R^{n\times n}|\det(A)=1\},\cdot\rangle$ — мультипликативные неабелевы группы при $n>1$.

4) множество $S(X)$ всех биективных отображений $f\colon X\to X$ с операцией произведения (суперпозиции) отображений — мультипликативная неабелева группа при $|X|>2$ и абелева группа при $|X|\leqslant 2$. Частным случаем является множество биективных отображений конечного множества из n элементов в себя. Это группа называется симметрической группой n-го порядка.

Теорема 1.1.1 (Простейшие свойства группы).

Для любой группы $\langle\Gamma,\circ\rangle$ справедливы утверждения.

1. *Нейтральный элемент единственен.*

2. *Для любого элемента симметричный к нему элемент единственен.*

3. *Для любых элементов $x,y,z\in\Gamma$ из равенства $x\circ z=y\circ z$ или $z\circ x=z\circ y$ следует $x=y$ — сокращение обеих частей равенства на общий элемент.*

4. *Для любых элементов $a,b\in\Gamma$ уравнение $a\circ x=b$ имеет единственное решение $x=a'\circ b$ и уравнение $y\circ a=b$ имеет единственное решение $y=b\circ a'$, где a' — симметричный к a элемент. В частности,*

   для $\langle\Gamma,\cdot\rangle$ $x=a^{-1}b$, $y=ba^{-1}$, для $\langle\Gamma,+\rangle$ $x=(-a)+b$, $y=b+(-a)$.

5. *Для любых элементов $a,b\in\Gamma$ $(a\circ b)'=b'\circ a'$. В частности,*

   для $\langle\Gamma,\cdot\rangle$ $(ab)^{-1}=b^{-1}a^{-1}$, для $\langle\Gamma,+\rangle$ $-(a+b)=(-b)+(-a)$.

6. Для любого элемента $a \in \Gamma$ $(a')' = a$. *В частности,*

для $\langle \Gamma, \cdot \rangle$ $(a^{-1})^{-1} = a$, *для* $\langle \Gamma, + \rangle$ $-(-a) = a$.

Доказательство.

1. Если e_1 и e_2 — нейтральные элементы, то $e_1 \circ e_2 = e_2 \circ e_1 = e_2$, т. к. e_1 — нейтральный элемент и $e_1 \circ e_2 = e_2 \circ e_1 = e_1$, т. к. e_2 — нейтральный элемент. Отсюда получаем $e_1 \circ e_2 = e_2$ и $e_1 \circ e_2 = e_1 \Rightarrow e_1 = e_2$.

2. Пусть b_1 и b_2 — симметричные элементы к элементу a, т.е. $b_1 \circ a = a \circ b_1 = e$, $b_2 \circ a = a \circ b_2 = e$. Отсюда и из ассоциативности операции следует, что $(b_1 \circ a) \circ b_2 = b_1 \circ (a \circ b_2) \Rightarrow e \circ b_2 = b_1 \circ e \Rightarrow b_2 = b_1$.

3. Пусть $x \circ z = y \circ z$. Возьмём симметричный к z элемент z' и проведём над элементами обеих частей равенства $x \circ z = y \circ z$ справа операцию с z'. Получим $(x \circ z) \circ z' = (y \circ z) \circ z' \Rightarrow x \circ (z \circ z') = y \circ (z \circ z') \Rightarrow x \circ e = y \circ e \Rightarrow x = y$.

Аналогично из равенства $z \circ x = z \circ y$ получите $x = y$.

4. Проверим, что $x = a' \circ b$ решение уравнения $a \circ x = b$. Действительно, $a \circ a' \circ b = b \Leftrightarrow e \circ b = b \Leftrightarrow b = b$ — верно.

Пусть $x = d$ — решение уравнения $a \circ x = b$, тогда $a \circ d = b$. Использовав симметричный элемент a', получим равенства $a' \circ a \circ d = a' \circ b \Rightarrow e \circ d = a' \circ b \Rightarrow d = a' \circ b$. Следовательно, решение $a' \circ b$ единственно.

Аналогично докажите утверждение для уравнения $y \circ a = b$.

5. Пусть $f = a \circ b$, $g = b' \circ a'$. Проверим, что $f \circ g = g \circ f = e$. Действительно, $f \circ g = (a \circ b) \circ (b' \circ a') = a \circ (b \circ b') \circ a' = a \circ e \circ a' = a \circ a' = e$. Аналогично, $g \circ f = (b' \circ a') \circ (a \circ b) = b' \circ (a' \circ a) \circ b = b' \circ e \circ b = b' \circ b = e$.

6. Из аксиомы группы (АГ3) следуют равенства $a \circ a' = a' \circ a = e$. Тогда симметричным для a' является элемент a, т. е. $(a')' = a$. ■

Задача 1.1.1. *Докажите утверждение: если* $\langle \Gamma, \cdot \rangle$ *— группа, то* $\forall\, n \in N\, (n \geqslant 3)$, $a_1, a_2, \ldots, a_n \in \Gamma$ *результат вычисления* $a_1 a_2 \ldots a_n$ *не зависит от расстановки скобок.*

Р е ш е н и е. Будем считать, что при вычислении выражения $a_1 a_2 \ldots a_n$ имеются обязательные $n-2$ пары открывающихся и закрывающихся скобок, фиксирующих нужный порядок вычисления результата

$n-1$ операции, и могут быть пары открывающихся и закрывающихся скобок, которые не влияют на результат и которые можно удалить.

Утверждение докажем ММИ по n.

1) При $n=3$ равенство $(a_1a_2)a_3 = a_1(a_2a_3)$ верно по (АГ1).

2) Предположим, что утверждение верно при всех $n<p$, где $p \in N$, $p \geqslant 4$, и докажем утверждение при $n=p$.

Для этого докажем, что результат вычисления $a_1a_2\ldots a_p$ при произвольной расстановке скобок даёт такой же результат, как при левой расстановке скобок: $(\ldots((a_1a_2)a_3)\ldots a_{p-1})a_p$, где все $p-2$ открывающиеся скобки стоят перед a_1, а закрывающиеся скобки по одной стоят после каждого элемента $a_2, a_3, \ldots, a_{p-1}$. Заметим, что при $p=3$ это выполнено: $(a_1a_2)a_3 = a_1(a_2a_3)$.

Обозначим через A результат вычисления $a_1a_2\ldots a_p$ при произвольной расстановке скобок и выделим последнюю операцию $A=(B)(C)$. Здесь возможны два случая 1) (C) совпадает с a_p и 2) C содержит операции с $r>1$ последними элементами a_{p-r+1}, a_{p-r+2}, $\ldots$, a_p.

1) По предположению индукции для вычисления B можно сделать левую расстановку скобок: $A=(B)a_p=(\ldots((a_1a_2)a_3)\ldots a_{p-1})a_p$, т. е. A совпадает со значением при левой расстановке скобок.

2) По предположению индукции для вычисления C можно сделать левую расстановку скобок:
$A=(B)(C)=(B)((\ldots(a_{p-r+1}a_{p-r+2})\ldots a_{p-1})a_p)$. Отсюда, обозначив $D=(\ldots(a_{p-r+1}a_{p-r+2})\ldots)a_{p-1}$, имеем $A=(B)((D)a_p)$. Далее после перестановки скобок для трёх элементов приходим к равенству $A = ((B)(D))a_p$. Здесь по предположению индукции для вычислении выражения $(B)(D)$ можно сделать левую расстановку скобок и как в п. 1) получим $A=(\ldots((a_1a_2)a_3)\ldots a_{p-1})a_p$, т. е. A совпадает со значением при левой расстановке скобок.

1.2. Подгруппа. Критерий подгруппы. Пересечение подгрупп

Будем использовать терминологию и обозначения для мультипликативных групп.

Определение 1. *Подмножество H группы Γ называется её подгруппой, если оно само является группой относительно алгебраической операции в Γ. Итак, если $\langle \Gamma, \cdot \rangle$ — группа, $H \subseteq \Gamma$ и $\langle H, \cdot \rangle$ — группа, то H является подгруппой группы Γ.*■

Теорема 1.2.1 (Критерий подгруппы). *Пусть $\langle \Gamma, \cdot \rangle$ — группа. Подмножество $H \neq \emptyset$ группы Γ является подгруппой этой группы тогда и только тогда, когда*

(КПГ1) $\quad \forall a, b \in H \quad ab \in H$;

(КПГ2) $\quad \forall a \in H \quad a^{-1} \in H$.

Доказательство.

1) Пусть H является подгруппой группы $\langle \Gamma, \cdot \rangle$. По определению $\langle H, \cdot \rangle$ является группой, следовательно, в H определена операция: $\forall\, a, b \in H\, \exists!\, w \in H \,|w = ab$ и выполнена аксиома (АГ3): $\forall\, a \in H\, \exists\, a^{-1} \in H \quad aa^{-1} = a^{-1}a = e$, т. е. выполнены (КПГ1) и (КПГ2).

2) Пусть $H \neq \emptyset$ и выполнены (КПГ1), (КПГ2).

I. Из (КПГ1) следует, что для любых элементов $a, b \in H$ операция ab определена в H.

II. Проверим, что в H выполнены аксиомы группы.

Возьмём любые $a, b, c \in H$. Поскольку $H \subseteq \Gamma$ и в Γ выполнена аксиома (АГ1), то выполнено равенство $(ab)c = a(bc)$. Из (КПГ1) следует, что элементы ab, $(ab)c$, bc, $a(bc)$ принадлежат H, поэтому в H выполнена аксиома (АГ1).

Покажем, что единица e группы Γ принадлежит H.

Возьмём любой элемент $a \in H$. Из (КПГ2) следует, что обратный элемент $a^{-1} \in H$. Затем из (КПГ1) получаем, что $aa^{-1} \in H$, и по аксиоме (АГ3) в группе Γ имеем $aa^{-1} = e$. Следовательно, $e \in H$.

Аксиомы (АГ2), (АГ3) выполнены в H, так как $\forall\, a \in H \subseteq \Gamma$ выполнены равенства $ae = ea = a$, $aa^{-1} = a^{-1}a = e$, где a, a^{-1}, $e \in H$. ■

С помощью критерия подгруппы легко обосновать следующие примеры.

Примеры.

1. $\langle Z, +\rangle$, $\langle Q, +\rangle$ — подгруппы группы $\langle R, +\rangle$.

2. Для любого числа $m \in N$ обозначим через mZ все целые числа кратные m. Тогда $\langle mZ, +\rangle$ — подгруппа группы $\langle Q, +\rangle$, а в силу примера 1 также подгруппа групп $\langle Z, +\rangle$, $\langle Q, +\rangle$, $\langle R, +\rangle$.

Теорема 1.2.2 (Пересечение подгрупп).

Пусть $\langle \Gamma, \cdot\rangle$ — группа и H_α — семейство её подгрупп, зависящих от индекса α, который пробегает множество A. Тогда пересечение всех подгрупп $\bigcap\limits_{\alpha\in A} H_\alpha$ является подгруппой группы Γ.

Доказательство. Обозначим $F = \bigcap\limits_{\alpha\in A} H_\alpha$.

Поскольку все подгруппы содержит единицу e группы Γ, то и их пересечение F содержат e, т. е. F не пусто.

Для доказательства утверждения применим критерий подгруппы.

(КПГ1): $a, b \in F \Rightarrow ab \in F$.

Пусть $a, b \in F = \bigcap\limits_{\alpha\in A} H_\alpha \Rightarrow \forall\alpha \in A \;\; a, b \in H_\alpha \Rightarrow \forall\alpha \in A \;\; ab \in H_\alpha \Rightarrow ab \in \bigcap\limits_{\alpha\in A} H_\alpha = F \Rightarrow$ выполнено (КПГ1).

(КПГ2): $a \in F \Rightarrow a^{-1} \in F$.

Пусть $a \in F = \bigcap\limits_{\alpha\in A} H_\alpha \Rightarrow \forall\alpha \in A \;\; a \in H_\alpha \Rightarrow \forall\alpha \in A \;\; a^{-1} \in H_\alpha \Rightarrow a^{-1} \in \bigcap\limits_{\alpha\in A} H_\alpha = F \Rightarrow$ выполнено (КПГ2).

Следовательно, H — подгруппа. ■

1.3. Произведение (сумма) подмножеств группы. Произведение (сумма) подгрупп

Определение 1. *Пусть S и T — непустые подмножества группы Γ.*

Для мультипликативной группы произведением подмножеств S и T называется множество $\{z|z \in \Gamma,\ z = ab,\ a \in S,\ b \in T\}$. Это множество обозначается ST. Операция $\langle S, T\rangle \mapsto ST$ называется умножением подмножеств.

Для аддитивной группы суммой подмножеств S и T называется множество $\{z|z \in \Gamma,\ z = a + b,\ a \in S,\ b \in T\}$. Это множество обозначается $S + T$. Операция $\langle S, T\rangle \mapsto S + T$ называется сложением подмножеств. ■

Теорема 1.3.1. 1° *На непустых подмножествах группы операция умножения (сложения) ассоциативна, т. е. для любых непустых подмножеств S, T, U группы выполнено равенство $(ST)U = S(TU)$ $\big((S + T) + U = S + (T + U)\big)$.*

2° *На непустых подмножествах коммутативной группы операция умножения (сложения) коммутативна, т. е. для любых непустых подмножеств S, T группы выполнено равенство $ST = TS$ $\big(S + T = T + S\big)$.*

Доказательство.

1° $ST = \{ab,\ a \in S,\ b \in T\}$, $(ST)U = \{(ab)c \,|\, a \in S,\ b \in T, c \in U\}$,
$TU = \{bc \,|\, b \in T,\ c \in U\}$, $S(TU) = \{a(bc) \,|\, a \in S,\ b \in T,\ c \in C\}$.

Из $(ab)c = a(bc)$ следует, что $(ST)U = S(TU)$.

2° $ST = \{ab,\ a \in S,\ b \in T\} = \{ba,\ a \in S,\ b \in T\} = TS$.

Очевидно, что замена операции умножения на сложение не меняет сути доказательства. ■

Замечания. 1. В дальнейшем для групп будем использовать мультипликативную терминологию и обозначения.

2. Для ST будем писать а) aT вместо $\{a\}T$, если $S = \{a\}$ и T содержит более одного элемента, б) Sb вместо $S\{b\}$, если S содержит более одного элемента и $T = \{b\}$, в) ab вместо $\{a\}\{b\}$, если $S = \{a\}$ и $T = \{b\}$. ■

Теорема 1.3.2. *Если H_1, H_2 — подгруппы группы Γ и $H_1H_2 = H_2H_1$, то произведение H_1H_2 является подгруппой группы Γ.*

Доказательство. Обозначим $H = H_1H_2 = H_2H_1$ и применим к H критерий подгруппы.

(КПГ1): $a, b \in H \Rightarrow ab \in H$.

$a, b \in H = H_1H_2 \Rightarrow (a = a_1a_2,\ a_1 \in H_1,\ a_2 \in H_2) \bigwedge (b = b_1b_2,\ b_1 \in H_1,\ b_2 \in H_2) \Rightarrow ab = (a_1a_2)(b_1b_2)$ {используем ассоциативность} $\Rightarrow ab = a_1(a_2b_1)b_2$.
Здесь $a_2b_1 \in H_2H_1 = H_1H_2 \Rightarrow a_2b_1 \in H_1H_2 \Rightarrow a_2b_1 = d_1d_2,\ d_1 \in H_1,\ d_2 \in H_2$.
Тогда $ab = a_1(d_1d_2)b_2$ {используем ассоциативность} $\Rightarrow ab = (a_1d_1)(d_2b_2)$, {где $a_1d_1 \in H_1,\ d_2b_2 \in H_2$} $\Rightarrow ab \in H_1H_2 = H \Rightarrow ab \in H$.

(КПГ2): $a \in H \Rightarrow a^{-1} \in H$.

$a \in H = H_1H_2 \Rightarrow a = a_1a_2,\ a_1 \in H_1,\ a_2 \in H_2 \Rightarrow a^{-1} = a_2^{-1}a_1^{-1},\ a_2^{-1} \in H_2,\ a_1^{-1} \in H_1 \Rightarrow a^{-1} \in H_2H_1 = H \Rightarrow a^{-1} \in H$. ■

Следствие. *Если H_1, H_2 — подгруппы коммутативной группы Γ, то произведение H_1H_2 является подгруппой группы Γ.*

Задача 1.3.1. *Если $H_1,\ H_2, \ldots,\ H_n$ подгруппы группы $\langle \Gamma, \cdot \rangle$ и $H_iH_j = H_jH_i$ при всех $i \neq j$, то произведение $H_1H_2 \ldots H_n$ является подгруппой группы Γ.*

Задача 1.3.2. *Проверить является ли данное множество группой или абелевой группой: $\langle R, + \rangle$.*

Р е ш е н и е.

Будем использовать без доказательства следующие свойства на множестве $M = R,\ Q,\ Z,\ N$:

(С1) $\forall\, a, b, c \in M\ (a + b) + c = a + (b + c)$ — операция сложения ассоциативна;

(С2) $\forall\, a, b \in M \quad a + b = b + a$ — операция сложения коммутативна;

(С3) $\forall\, a, b, c \in M \quad (ab)c = a(bc)$ — операция умножения ассоциативна;

(С4) $\forall\, a, b \in M \quad ab = ba$ — операция умножения коммутативна;

(С5) $\forall\, a, b, c \in M \quad (a + b)c = ac + bc, \quad c(a + b) = ca + cb$ — операция умножения дистрибутивна относительно сложения справа и слева.

Проверяем по определению, является ли $\langle R, + \rangle$ группой.

I. На R определена операция сложения: $\forall a, b \in R \ \ a + b \in R$.

II. (АГ1) $\forall\, a, b, c \in R\ (a+b)+c = a+(b+c)$ — верно для чисел из R.

(АГ2) $\exists\, e \in R\, \forall\, a \in R \quad a+e = e+a = a$. Здесь мы ищем такое $e \in R$, для которого при любых $a \in R$ выполнены равенства $a+e = e+a = a$. Это верно при $e = 0$.

(АГ3) изложим в других обозначениях: $\forall\, a \in R\, \exists\, x \in R \quad a+x = x+a = 0$. Здесь для каждого $a \in R$ ищем такое $x \in R$ (неизвестное), для которого выполнены равенства $a+x = x+a = 0$. Это верно при $x = (-1)a$.

Из (АГ1) – (АГ3) $\Rightarrow \langle R, +\rangle$ — группа.

(АГ4) $\forall\, a, b \in R \quad a+b = b+a$ верно для чисел из R.

Из (АГ1) – (АГ4) $\Rightarrow \langle R, +\rangle$ — абелева группа.

Задача 1.3.3. *Проверить является ли данное множество группой или абелевой группой:* $\langle Q, +\rangle$.

Р е ш е н и е. В задаче 1.3.2 доказано, что $\langle R, +\rangle$ — абелева группа. Множество рациональных чисел Q является подмножеством множества вещественных чисел R. Для проверки является ли $\langle Q, +\rangle$ группой, воспользуемся критерием подгруппы.

(КПГ1) $a, b \in Q \Rightarrow a+b \in Q$ выполнено, т. к. если $a = \dfrac{m}{n}$, $b = \dfrac{p}{q}$, где $m, p \in Z$, $n, q \in N$, то $a+b = \dfrac{mq+np}{nq} \in Q$.

(КПГ2) $a \in Q \Rightarrow (-a) \in Q$ выполнено, т. к. если $a = \dfrac{m}{n}$, где $m \in Z$, $n \in N$, то $(-a) = \dfrac{-m}{n} \in Q$.

Из (КПГ1) – (КПГ2) $\Rightarrow$ $\langle Q, +\rangle$ — подгруппа группы $\langle R, +\rangle$ $\Rightarrow$ $\langle Q, +\rangle$ — группа. Из того, что $\langle R, +\rangle$ — абелева группа и выполнено (КПГ1) следует, что $\langle Q, +\rangle$ — абелева группа.

Задача 1.3.4. *Проверить является ли группой или абелевой группой каждое из следующих множеств:*
1) $\langle N, +\rangle$; 2) $\langle R^{m\times n}, +\rangle$; 3) $\langle N, \cdot\rangle$; 4) $\langle Z, \cdot\rangle$; 5) $\langle R \setminus 0, \cdot\rangle$; $\langle Q \setminus 0, \cdot\rangle$; $\langle \{A \in R^{n\times n} | \det(A) \neq 0\}, \cdot\rangle$, $\langle \{A \in R^{n\times n} | \det(A) = 1\}, \cdot\rangle$ — *мультипликативные неабелевы группы при* $n > 1$.

множество $S(X)$ всех биективных отображений $f\colon X \to X$ с операцией произведения (суперпозиции) отображений — мультипликативная неабелева группа. Частным случаем является множество биективных отображений конечного множества из n элементов в себя. Это группа называется симметрической группой n-го порядка.

Решения или ответы.

1) I. На N определена операция сложения: $\forall a, b \in N \ a + b \in N$.

II. (АГ1) $\forall\, a, b, c \in N \ (a + b) + c = a + (b + c)$ — верно для чисел из N.

(АГ2) $\exists\, e \in N \, \forall\, a \in N \quad a + e = e + a = a$ не выполнена при любом $e \in N$, т. к. если $e \in N$, то $a + e > a \ \Rightarrow \langle N, + \rangle$ не является группой.

2) I. На Z определена операция умножения: $\forall a, b \in Z \ \ ab \in Z$.

II. (АГ1) $\forall\, a, b, c \in Z \ (ab)c = a(bc)$ — верно для чисел из Z.

(АГ2) $\exists\, e \in Z \, \forall\, Z \in R \quad ae = ea = a$. Это верно при $e = 1$.

(АГ3) $\forall\, a \in Z \, \exists\, x \in Z \quad ax = xa = 1$.

Это неверно для чисел $a \neq \pm 1$. Самый простой пример — $a = 0$. Уравнение $0x = 1$ не имеет решений.

Другой пример — $a = 2$. Уравнение $2x = 1$ для вещественных x имеет единственное решение $x = \dfrac{1}{2}$. Поэтому уравнение $2x = 1$ не имеет целых решений.

Итак, (АГ2) неверна и $\langle Z, \cdot \rangle$ не является группой.

1.4. Смежные классы группы и их свойства

Определение 1. *Пусть H — подгруппа группы Γ, a — элемент группы Γ. Множество aH (множество Ha) называется левым (правым) смежным классом группы Γ по подгруппе H, порождённым элементом a.* ■

Теорема 1.4.1 (Простейшие свойства смежных классов). *Пусть H — подгруппа группы Γ. Тогда выполнены следующие свойства.*

1° $\forall\, a \in \Gamma \ \ a \in aH$ *и* $a \in Ha$.

2° *Левый (правый) смежный класс состоит из элементов группы, причём любой элемент группы входит в некоторый левый и в некоторый правый смежный класс.*

3° *Подгруппа H является левым и правым смежным классом.*

4° *В коммутативной группе $aH = Ha \ \ \forall\, a \in \Gamma$.*

5° *Смежный класс порождается любым своим элементом.*

6° *Любые два левых (правых) смежных класса либо совпадают, либо не пересекаются.*

Доказательство.

1° Поскольку $e \in H$, имеем $a = ae \in aH$ и $a = ea \in Ha$.

2° По определению для $a \in \Gamma$ имеем $aH \subseteq \Gamma$ и $Ha \subseteq \Gamma$. Из 1° следует, что $a \in aH$ и $a \in Ha$.

3° $H = eH = He$.

4° В коммутативной группе Γ для всех элементов выполнено равенство $ab = ba$, поэтому $aH = Ha \ \ \forall\, a \in \Gamma$.

5° Для левого смежного класса aH докажем, что если $b \in aH$, то $aH = bH$, т. е. класс aH порождается любым элементом $b \in aH$.

Из $b \in aH$ следует, что $b = ah_1$, где $h_1 \in H$.

I. Проверим, что $bH \subseteq aH$.

Берём любой элемент $z \in bH \Rightarrow z = bh_2$, где $h_2 \in H$ и $b = ah_1 \Rightarrow$ $\Rightarrow z = ah_1h_2$, где $h_1h_2 \in H \Rightarrow z \in aH$. Следовательно, $bH \subseteq aH$.

II. Проверим, что $aH \subseteq bH$.

Берём любой элемент $z \in aH \Rightarrow z = ah_3$, где $h_3 \in H$. Далее из равенства $b = ah_1$ выразим $a = bh_1^{-1}$, где $h_1^{-1} \in H$. Тогда $z = ah_3 = bh_1^{-1}h_3$, где $h_1^{-1}h_3 \in H$. Следовательно, $z \in bH$ и $aH \subseteq bH$.

Из $bH \subseteq aH$ и $aH \subseteq bH$ следует равенство $aH = bH$.

Докажите сами, что при $b \in Ha$ выполнено равенство $Ha = Hb$.

6° Из утверждения 5° следует, что если два левых смежных класса aH и bH имеют общий элемент z, то они совпадают, т. к. $zH = aH$, $zH = bH$. А если левые смежные классы не имеют общих элементов, то они не пересекаются. Такое же рассуждение верно и для правых смежных классов.

Следствие. Группа разбивается на непересекающиеся левые (правые) смежные классы по заданной подгруппе. Это разбиение называется *левосторонним (правосторонним) разложением группы по подгруппе.* ■

Примеры смежных классов.

Группа	Подгруппа	Смежные классы
$\langle Z, +\rangle$	$H = mZ = \{mk \mid k \in Z\}$, $m \in N, m > 1$	Для $a \in Z : a + H = H + a = \{n \mid n = a + mk,\ k \in Z\}$. Здесь m различных классов: $a + H$, $a = 0, 1, \ldots, m - 1$. Числа из такого класса $a + H$ при делении на m дают остаток a.

1.5. Конечные группы и их построение. Теорема Лагранжа о порядке подгруппы

Определение 1. *Группа* Γ, *состоящая из конечного числа элементов, называется конечной группой. Число элементов этой группы называется порядком группы и обозначается* $\mathrm{card}(\Gamma)$, $|\Gamma|$, $\overline{\overline{\Gamma}}$. ■

Алгебраическая операция в конечной группе может быть представлена «таблицей умножения».

1. Построим группу 1-го порядка. Возьмём множество $\Gamma = \{e\}$ из одного элемента, который назовём единицей. Операцию определим равенством $ee = e$ или таблицей на рис. 1.

Выполнение аксиом очевидно, АГ1: $(ee)e = e(ee)$ (ассоциативность); АГ2: $ee = ee = e$ (единица — это e); АГ3: $ee = ee = e$ (обратный к e — это e); АГ4: $ee = ee$ (коммутативность).

2. Построим группу 2-го порядка.

Возьмём множество из двух элементов $\Gamma = \{e, a\}$, e — единица. Определим операцию как в таблице на рис. 2: 1) $ee = e$, 2) $ea = ae = a$, 3) $aa = x$ и выберем x. Для группы в каждой строке (каждом столбце) таблицы все элементы различны, поскольку из равенства $fg = fh\,(gf = hf)$ после сокращения получаем $g = h$, что не верно. Значит, $x = e$ (см. таблицу на рис. 3).

Проверим аксиомы группы.

1-й \ 2-й	e
e	e

Рис. 1

1-й \ 2-й	e	a
e	e	a
a	a	x

Рис. 2

1-й \ 2-й	e	a
e	e	a
a	a	e

Рис. 3

Из симметрии таблицы следует коммутативность операции. Равенства $ee = e$, $ea = ae = a$ показывают, что единица — это e. Равенства $ee = e$, $aa = e$ обеспечивает наличие обратных элементов для e и a.

Проверим ассоциативность операции: $\forall x, y, z \in \Gamma \quad (xy)z = x(yz)$.

I. Пусть среди трёх элементов хотя бы один равен e :

а) 1-й e, 2-й k, 3-й m, тогда $(ek)m = km$ и $e(km) = km$, т.е. $(ek)m = e(km)$;

б) 1-й k, 2-й e, 3-й m, тогда $(ke)m = km$ и $k(em) = km$, т.е. $(ke)m = k(em)$;

в) 1-й k, 2-й m, 3-й e, тогда $(km)e = km$ и $k(me) = km$, т.е. $(km)e = k(me)$.

II. Пусть среди трёх элементов ни один не равен e.

Тогда все три элемента равны a, и в силу коммутативности $(aa)a = a(aa)$.

Итак, аксиомы абелевой группы выполнены.

3. Построим группу 3-го порядка. Возьмём множество $\Gamma = \{e, a, b\}$, e — единица. Определим операцию как в таблице на рис. 4: 1) $ee = e$, 2) $ea = ae = a$, 3) $eb = be = b$, 4) $aa = x$, 5) $ab = y$, 6) $ba = z$, 7) $bb = u$ и выберем x, y, z, u. Равенства 1) – 3) обеспечивают выполнение АГ2. Для группы в каждой строке (каждом столбце) таблицы все элементы

1-й \ 2-й	e	a	b
e	e	a	b
a	a	x	y
b	b	z	u

Рис. 4

1-й \ 2-й	e	a	b
e	e	a	b
a	a	b	e
b	b	z	u

Рис. 5

1-й \ 2-й	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Рис. 6

различны. Во 2-й строке $x \neq a$, $y \neq a$, в 3-м столбце $y \neq b$, тогда $x = b$, $y = e$ Далее во 2-м и 3-м столбцах $z = e$, $u = a$ (см. таблицу на рис.

6). Из симметрии таблицы следует АГ4. АГ3 выполнена, т. к. обратные к e — это e, к a — это b, к b — это a.

Проверим АГ1: $\forall x, y, z \in \Gamma \quad (xy)z = x(yz)$.

I. Если среди трёх элементов хотя бы один равен e, то рассуждаем как в примере 2.

II. Пусть среди трёх элементов ни один не равен e.

1) Три a или три b : $(aa)a = a(aa)$ и $(bb)b = b(bb)$ в силу АГ4.

2) Два a и один b :

$(aa)b = a(ab)$, т. к. $(aa)b = bb = a$ и $a(ab) = ae = a$;

$(ab)a = a(ba)$, т. к. $(ab)a = ea = a$ и $a(ba) = ae = a$;

$(ba)a = b(aa)$, т. к. $(ba)a = ea = a$ и $b(aa) = bb = a$.

2) Один a и два b :

$(ab)b = a(bb)$, т. к. $(ab)b = eb = b$ и $a(bb) = aa = b$;

$(ba)b = b(ab)$, т. к. $(ba)b = eb = b$ и $b(ab) = be = b$;

$(bb)a = b(ba)$, т. к. $(bb)a = aa = b$ и $b(ba) = be = b$.

Итак, Γ — абелева группа. ■

Теорема 1.5.1 (Лагранж). *Порядок подгруппы является делителем порядка конечной группы.*

Доказательство. Пусть $H = \{h_1, h_2, \ldots, h_k\}$ — подгруппа группы $\Gamma = \{g_1, g_2, \ldots, g_n\}$, причём все элементы различны: $h_p \neq h_q$ при $p \neq q$ и $g_i \neq g_j$ при $i \neq j$. Нужно доказать, что n делится на k. Рассмотрим левостороннее разложение группы Γ по подгруппе H. Оно состоит из конечного числа (обозначим его m) непересекающихся левых смежных классов вида a_iH, где $a_i \in \Gamma$ — порождающий элемент класса, $i = \overline{1, m}$, причём $m \leqslant n$. Имеем $\Gamma = \bigcup\limits_{i=1}^{m} a_iH$ и количество элементов группы равно сумме числа элементов указанных классов. Покажем что все k элементов любого класса $a_iH = \{a_ih_1, a_ih_2, \ldots, a_ih_k\}$ различны. Действительно, из равенства $a_ih_p = a_ih_q$ при $p \neq q$ после сокращения получаем $h_p = h_q$, но $h_p \neq h_q$. Таким образом, количество всех элементов группы равно mk, т. е. $n = mk$. ■

1.6. Нормальный делитель. Фактор-группа

Определение 1. *Подгруппа H группы Γ называется нормальным делителем, если $\forall g \in \Gamma$ $gH = Hg$, т. е. для каждого элемента группы его левый и правый смежные классы по подгруппе H совпадают.* ■▌

В абелевой группе любая подгруппа H является нормальным делителем, т. к. $\forall g \in \Gamma$ $gH = Hg$ в силу коммутативности операции.

Теорема 1.6.1 (Критерий нормального делителя).
Подгруппа H группы Γ является нормальным делителем тогда и только тогда, когда $g^{-1}hg \in H$ при всех $g \in \Gamma$ и $h \in H$.

Доказательство. I. Пусть H — нормальный делитель группы Γ, т. е. для произвольного $g \in \Gamma$ $gH = Hg$. Отсюда следует включение $Hg \subseteq gH$. Оно означает, что любой элемент из Hg принадлежит gH, т. е. $\forall h \in H \, \exists \, h_1 \in H$ такое, что $hg = gh_1$. Умножив слева на обратный g^{-1}, получим $g^{-1}hg = h_1$, где $h_1 \in H$. Следовательно, $\forall \, g \in \Gamma, h \in H \colon g^{-1}hg \in H$.

II. Пусть $\forall \, g \in \Gamma, h \in H$ $g^{-1}hg \in H$. Возьмём любые элементы $h \in H$ и $g \in \Gamma$, тогда $g^{-1}hg \in H$ и $ghg^{-1} = (g^{-1})^{-1}hg^{-1} \in H$. Обозначим элементы $g^{-1}hg = h_1$, $ghg^{-1} = h_2$, причём $h_1, h_2 \in H$. Отсюда $hg = gh_1$, $gh = h_2g$. Тогда $Hg \subseteq gH$ и $gH \subseteq Hg$, т. е. $gH = Hg$.

Теорема 1.6.2 (Построение фактор-группы).
Смежные классы по нормальному делителю образуют группу относительно умножения подмножеств группы.

Доказательство.
Пусть подгруппа H группы Γ является нормальным делителем.
а) Сначала покажем, что $HH = H$.
Во-первых, $\forall a, b \in H$ по критерию подгруппы $ab \in H$, т. е. $HH \subseteq H$.
Во-вторых, $\forall a \in H$ $a = ae$, где $e \in H$, тогда $ae \in HH$, т. е. $H \subseteq HH$.
б) Проверим, что произведение левых смежных классов aH и bH является левым смежным классом $(ab)H$:

$$(aH)(bH) = \left\{ \begin{array}{c} \text{ассоциативность} \\ \text{умножения} \end{array} \right\} = a(Hb)H = \left\{ \begin{array}{c} Hb = bH, \text{ т. к.} \\ H \text{ — нормальный} \\ \text{делитель} \end{array} \right\} =$$

$$= a(bH)H = \left\{ \begin{array}{c} \text{ассоциативность} \\ \text{умножения} \end{array} \right\} = (ab)(HH) = \{\, HH = H \,\} = (ab)H.$$

в) Ассоциативность умножения классов:

$((aH)(bH))(cH) = ((ab)H)(cH) = (abc)H = (aH)((bc)H) = (aH)((bH)(cH))$. ▮

г) Единица для умножения классов $eH = H$:

$(aH)(eH) = (ae)H = aH$, $(eH)(aH) = (ea)H = aH$.

д) Обратный классу aH — это класс $a^{-1}H$:

$(aH)(a^{-1}H) = (aa^{-1})H = eH = H$, $(a^{-1}H)(aH) = (a^{-1}a)H = eH = H$. ■

Замечание. В аддитивной группе Γ с нормальным делителем H, смежные классы $a+H$ для $a \in \Gamma$ образуют фактор-группу. В ней операция сложения задана равенством $(a+H)+(b+H) = (a+b)+H$, нулевой элемент равен $0+H = H$, противоположный для $a+H$ — это $(-a)+H$.

Определение 2. *Группа смежных классов группы* Γ *по нормальному делителю* H *называется* **фактор-группой группы Γ по подгруппе H**.

Обозначение: $\Gamma|H$.

В примере на стр. 17 в аддитивной абелевой группе Z указана подгруппа mZ, которая является нормальным делителем, и приведены смежные классы группы Z по подгруппе mZ — это множества $a+mZ = \{n = a + mk,\ k \in Z\}$ для $a \in Z$. Всего имеется m различных смежных классов: $a + mZ$, $a = \overline{0, m-1}$. Эти смежные классы образуют фактор-группу $Z|mZ$.

1.7. Кольцо. Простейшие свойства кольца. Подкольцо. Критерий подкольца. ▮

Определение 1. *Непустое множество* K *называется кольцом, если*

(I) *на множестве* K *определены две алгебраические операции*

а) *сложения* $\forall\, a, b \in K\, \exists!\, z \in \Gamma$, *обозначаемый* $z = a + b$,

б) *умножения* $\forall\, a, b \in K\, \exists!\, w \in K$, *обозначаемый* $w = a \cdot b$ *или* $w = ab$,

(II) *выполнены следующие аксиомы*

(AK1) $\forall\, a, b, c \in K\ (a + b) + c = a + (b + c)$ — *операция сложения ассоциативна;*

(AK2) $\exists\, 0 \in K\, \forall\, a \in K \quad a + 0 = a$ — *операция сложения имеет нейтральный элемент* 0 *(нуль);*

(AK3) $\forall\, a \in K\, \exists\, b \in K \quad a + b = 0$ — *для каждого элемента* $a \in K$ *существует противоположный элемент* $b \in K$, *обозначаемый* $(-a)$;

(AK4) $\forall\, a, b \in K \quad a + b = b + a$ — *операция сложения коммутативна;*

(AK5) $\forall\, a, b, c \in K \quad (ab)c = a(bc)$ — *операция умножения ассоциативна;*

(AK6) $\forall\, a, b, c \in K \quad (a + b)c = ac + bc,\ c(a + b) = ca + cb$ — *операция умножения дистрибутивна относительно сложения справа и слева.*

Обозначение: кольцо K *или* $\langle K, +, \cdot \rangle$.

Группу $\langle K, + \rangle$ *называют аддитивной группой кольца.*

($\star$) *Если в* K *существует нейтральный элемент (единица) для умножения, т. е.* $\exists e \in K\, \forall a \in K \quad ae = ea = a$, *то* K *называется кольцом с единицей.*

($\diamond$) *Если в* K *операция умножения коммутативна, т. е.* $\forall a, b \in K \quad ab = ba$, *то* K *называется коммутативным кольцом.*

($\triangleright$) *Если в* K *существуют ненулевые элементы* a *и* b *такие, что* $ab = 0$, *то они называются делителями нуля* (a — *левый,* b — *правый делитель*), а кольцо называется кольцом с делителями нуля. ■

Замечания.

1. Из АК2 и АК4 следует, что $\exists\theta \in P\colon \forall a \in P \quad a+\theta=\theta+a=a$.
2. Из АК3 и АК4 следует, что $\forall a \in P\ \exists(-a) \in P\colon a+(-a)=(-a)+a=\theta$.

Примеры. Множество квадратных матриц $R^{n\times n}$ — кольцо с единицей, некоммутативное при $n>1$; Z, Q, R — коммутативные кольца с единицей; $2Z$ — коммутативное кольцо без единицы.

Теорема 1.7.1 (Простейшие свойства кольца).

Для любого кольца K справедливы утверждения.

1. *Кольцо обладает всеми свойствами аддитивной абелевой группы, в частности, в кольце:*
   - (a) *нуль единственен;*
   - (b) *для любого элемента противоположный к нему элемент единственен;*
   - (c) *из условия $a,b,c \in K$ и $a+c=b+c$ или $c+a=c+b$ следует равенство $a=b$;*
   - (d) *для любых элементов $a,b \in K$ существует, и притом единственное, решение $x=b+(-a)$ уравнения $x+a=b$; это решение называется* **разностью** *b и a и обозначается $b-a$. Итак, $b-a=b+(-a)$ и эта формула определяет операцию* **вычитания**.
2. *Умножение дистрибутивно относительно вычитания, т.е.*

   $\forall a,b,c \in K \quad (a-b)c=ac-bc$ *и* $c(a-b)=ca-cb$.
3. $\forall a \in K \quad a0=0$ *и* $0a=0$.
4. *В кольце с единицей она единственна (и обычно обозначается 1).*
5. *В кольце K без делителей нуля из условия $a,b \in K$ и $ab=0$ следует условие $a=0$ или $b=0$.*
6. *В кольце K без делителей нуля из условий $a,b,c \in K$, $c \neq 0$ и $ac=bc$ или $ca=cb$ следует равенство $a=b$.*

7. *В кольце с единицей для любого элемента выполнены равенства* $(-a) = (-1)a = a(-1)$.

8. *В кольце для любых элементов выполнены равенства*

   а) $(-a)b = a(-b) = -(ab)$, *б)* $(-a)(-b) = ab$.

9. *В кольце с единицей, содержащем не менее двух элементов:* $1 \neq 0$.

Доказательство.

1a. Если 0_1 и 0_2 — нулевые элементы, то $0_1 + 0_2 = 0_1$, т.к. 0_2 — нулевой элемент и $0_2 + 0_1 = 0_2$ т.к. 0_1 — нулевой элемент. Отсюда и из коммутативности сложения $0_1 + 0_2 = 0_2 + 0_1$ получим $0_1 = 0_2$.

1b. Пусть b_1 и b_2— противоположные элементы к элементу a, т.е. $b_1 + a = 0$, $a + b_2 = 0$. Из ассоциативности сложения следует, что $(b_1 + a) + b_2 = b_1 + (a + b_2)$. Отсюда $0 + b_2 = b_1 + 0$ и $b_1 = b_2$.

1c. Прибавим противоположный элемент $(-c)$ к обеим частям равенства $a + c = b + c$. Получим $(a + c) + (-c) = (b + c) + (-c)$. Отсюда следуют равенства $a + (c + (-c)) = b + (c + (-c))$, $a + 0 = b + 0$ и $a = b$.

1d. Проверим, что $x = b + (-a)$ решение уравнения $x + a = b$. Действительно, $x + a = b + (-a) + a = b + 0 = b$.

Пусть $x = c_1$ и $x = c_2$ решения уравнения $x + a = b$. Тогда $c_1 + a = b$, $c_2 + a = b$ и $c_1 + a = c_2 + a$. Отсюда после сокращения $c_1 = c_2$, т.е. решение единственно.

2. Для разности $a - b$ справедливо равенство $(a - b) + b = a$. Умножим обе части этого равенства справа на c, получим $((a - b) + b)c = ac$. Далее из дистрибутивности умножения относительно сложения следует $(a - b)c + bc = ac$. Используя определение разности, получаем $(a - b)c = ac - bc$. Доказательство равенства $c(a - b) = ca - cb$ проведите самостоятельно.

3. Разность равных элементов равна 0 : $ab - ab = 0$, $a - a = 0$, $b - b = 0$. Отсюда, пользуясь 2, получим $a(b - b) = 0$, $(a - a)b = 0$. Следовательно, $a0 = 0$, $0b = 0$.

4. Если e_1 и e_2 — единицы, то $e_1e_2 = e_2e_1 = e_1$, т.к. e_2 — единица, и $e_2e_1 = e_1e_2 = e_2$, т.к. e_1 — единица. Выражения равны, следовательно,

$e_1 = e_2$.

5. $\sqcap$ $ab = 0$, где $a \neq 0$ и $b \neq 0$. Тогда a и b — делители нуля. $\sqcup$ Следовательно, $a = 0$ или $b = 0$.

6. Из условия $ac = bc$ и $c \neq 0 \Rightarrow ac - bc = 0 \Rightarrow (a-b)c = 0$, причём $c \neq 0$. Отсюда согласно свойству 5 имеем $a - b = 0$ и $a = b$. Аналогично, из $ca = cb$ и $c \neq 0 \Rightarrow c(a-b) = 0$, где $c \neq 0 \Rightarrow a - b = 0 \Rightarrow a = b$.

7. В 1b доказана единственность противоположного элемента. Сложим a и $(-1)a$, получим $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$. Отсюда $(-1)a$ является противоположным a, т. е. $(-1)a = (-a)$. Аналогично, $a + a(-1) = a1 + a(-1) = a(1 + (-1)) = a0 = 0 \Rightarrow a(-1) = (-a)$.

8a. $ab + (-a)b = (a + (-a))b = 0b = 0 \Rightarrow (-a)b = -(ab)$.

$ab + a(-b) = a(b + (-b)) = a0 = 0 \Rightarrow a(-b) = -(ab)$.

8b. Из $c = -(-c)$ следует, что $ab = -(-(ab))$. Отсюда, дважды используя 8a, получим $-(-(ab)) = -((-a)b) = (-a)(-b)$. Следовательно, $ab = (-a)(-b)$.

9. $\sqcap$ $1 = 0$. Тогда существует элемент $a \in K$ такой, что $a \neq 0$. Умножим равенство $1 = 0$ на a слева, получим $a \cdot 1 = a \cdot 0$ и $a = 0$. $\sqcup$

■

Определение 2. *Подмножество M кольца K называется* **подкольцом** *K, если M само является кольцом при тех же операциях сложения и умножения, которые заданы в кольце K.* ■

Теорема 1.7.2 (Критерий подкольца).

Пусть K — кольцо и M — непустое подмножество кольца K. M является подкольцом кольца K тогда и только тогда, когда M замкнуто относительно операций сложения, умножения и перехода к противоположному элементу, т. е.

(КПК1) *для любых элементов $a, b \in M$ $a + b \in M$;*

(КПК2) *для любых элементов $a, b \in M$ $ab \in M$;*

(КПК3) *для любого элемента $a \in M$ $(-a) \in M$.*

Доказательство.

1. Если M — подкольцо K, то M — кольцо и выполнены (КПК1) – (КПК3).

2. Пусть в M выполнены (КПК1) – (КПК3). Тогда в M определены операции сложения, умножения и любой элемент $a \in M$ имеет противоположный элемент $(-a) \in M$. Сложив эти элементы получим $a + (-a) \in M$, т.е. $0 \in M$. Тогда аксиомы кольца, выполненные в K, будут выполнены в M. Следовательно M — кольцо. ■

1.8. Поле

Определение 1. *Множество P называется полем, если на множестве P определены алгебраические операции сложения и умножения и выполнены следующие аксиомы*

АП1 $\forall a, b, c \in P \quad (a + b) + c = a + (b + c)$;

АП2 $\exists \theta \in P \, \forall a \in P \quad a + \theta = a$;

АП3 $\forall a \in P \, \exists b \in P \colon a + b = \theta$ *— для каждого элемента a существует противоположный элемент b, обозначаемый $(-a)$*;

АП4 $\forall a, b \in P \quad a + b = b + a$

АП5 $\forall a, b, c \in P \quad (ab)c = a(bc)$;

АП6 $\exists e \in P \setminus \theta \colon \forall a \in P \quad ae = a$;

АП7 $\forall a \in P \setminus \theta \, \exists b \in P \colon ab = e$ *— для каждого элемента $a \neq \theta$ существует обратный элемент b, обозначаемый a^{-1}*;

АП8 $\forall a, b \in P \quad ab = ba$;

АП9 $\forall a, b, c \in P \quad (a + b)c = ac + bc \Leftrightarrow c(a + b) = ca + cb$.

Замечания.

1. Из АП2 и АП4 следует, что $\exists \theta \in P \colon \forall a \in P \quad a + \theta = \theta + a = a$.
2. Из АП3 и АП4 следует, что $\forall a \in P \, \exists (-a) \in P \colon a + (-a) = (-a) + a = \theta$.
3. Из АП6 и АП8 следует, что $\exists e \in P \setminus \theta \colon \forall a \in P \quad ae = ea = a$.
4. Из АП7 и АП8 следует, что $\forall a \in P \setminus \theta \, \exists a^{-1} \in P \colon aa^{-1} = a^{-1}a = e$. ■

Примеры.

1. Множество рациональных чисел.
2. Множество действительных чисел.

Теорема 1.8.1 (Простейшие свойства поля).

1. *Нулевой элемент единственен и он обозначается* 0.

2. *Для любого элемента противоположный к нему элемент единственен.*

3. *Для любых элементов* $a, b, c \in P$ *из* $a + c = b + c$ *или* $c + a = c + b$ *следует* $a = b$.

4. *Для любых элементов* $a, b \in P$ *существует, и притом единственное, решение* $x = b+(-a)$ *уравнения* $x+a = b$; *это решение называется* **разностью** b *и* a *и обозначается* $b-a$. *Итак,* $b-a = b+(-a)$ *и эта формула определяет операцию* **вычитания**.

5. *Умножение дистрибутивно относительно вычитания, т.е.*

   $\forall a, b, c \in P \quad (a - b)c = ac - bc$ *и* $c(a - b) = ca - cb$.

6. $\forall a \in P \quad a0 = 0$ *и* $0a = 0$.

7. *Единичный элемент единственен (он обозначается* 1).

8. *Обратный элемент для любого ненулевого элемента единственен.*

9. *Для любых элементов* $a, b, c \in P$ *из* $c \neq 0$ *и* $ac = bc$ *или* $ca = cb$ *следует* $a = b$.

10. *Для любых элементов* $a, b \in P$ *из* $ab = 0$ *следует* $a = 0$ *или* $b = 0$.

11. $\forall a \in P \quad (-a) = (-1)a$ *(противоположный элемент для* a *равен* $(-1)a$).

12. $\forall a, b \in P \quad (-a)b = a(-b) = -(ab) = -ab$.

13. $\forall a, b \in P \quad (-a)(-b) = ab$.

14. *Для любых элементов* $a, b \in P$, $a \neq 0$ *существует, и притом единственное, решение* $x = ba^{-1}$ *уравнения* $xa = b$; *это решение* x *обозначается* $\dfrac{b}{a}$ *или* b/a *и называется* **частным** *или* **дробью**. *Итак,* $\dfrac{b}{a} = ba^{-1}$ *и эта формула определяет операцию* **деления**.

15. Справедливы следующие правила обращения с дробями:

(а) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b} = \dfrac{c}{d} \Leftrightarrow ad = bc;$

(б) $\forall a, b, c \in P,\ b \neq 0,\ c \neq 0 \quad \dfrac{a}{b} = \dfrac{ac}{bc};$

(в) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b}\dfrac{c}{d} = \dfrac{ac}{bd};$

(г) $\forall a, b, c \in P,\ b \neq 0 \quad \dfrac{a}{b} \pm \dfrac{c}{b} = \dfrac{a \pm c}{b};$

(д) $\forall a, b, c, d \in P,\ b \neq 0,\ d \neq 0 \quad \dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{ad \pm bc}{bd};$

(е) Для любых элементов $a, b \in P,\ b \neq 0 \quad \dfrac{-a}{b} = \dfrac{a}{-b} = -\dfrac{a}{b}.$

Доказательство.

1. Если θ_1 и θ_2 — нулевые элементы, то $\theta_1 + \theta_2 = \theta_1$, т.к. θ_2 — нулевой элемент, и $\theta_1 + \theta_2 = \theta_2$ т.к. θ_1 — нулевой элемент. Отсюда получим $\theta_1 = \theta_2$.

2. Если b_1 и b_2— противоположные элементы к a, то $b_1 + a = \theta$, $a + b_2 = \theta$. Из ассоциативности сложения следует, что $(b_1 + a) + b_2 = b_1 + (a + b_2)$. Отсюда $\theta + b_2 = b_1 + \theta$ и $b_1 = b_2$.

3. Прибавим противоположный элемент $(-c)$ к обеим частям равенства $a + c = b + c$. Получим $(a + c) + (-c) = (b + c) + (-c)$. Отсюда следуют равенства $a + (c + (-c)) = b + (c + (-c))$, $a + \theta = b + \theta$ и $a = b$.

4. Проверим, что $x = b + (-a)$ решение уравнения $x + a = b$. Действительно, $x + a = b + (-a) + a = b + \theta = b$.

Пусть $x = c$ решение уравнения $x + a = b$, т.е. $c + a = b$. Прибавим к обеим частям равенства $c + a = b$ противоположный элемент $(-a)$. Получим $c + a + (-a) = b + (-a)$. Откуда следует $c + \theta = b + (-a)$ и $c = b + (-a)$, т.е. решение единственно.

5. Для разности справедливо равенство $(a - b) + b = a$. Умножим обе части этого равенства справа на c, получим $((a - b) + b)c = ac$. Далее из дистрибутивности следует равенство $(a - b)c + bc = ac$. Используя сначала определение разности и затем коммутативность умножения, получаем $(a - b)c = ac - bc$ и $c(a - b) = ca - cb$.

6. Представим $0 = b - b$, $b \in P$. Умножим обе части этого равенства слева на a, получим $a0 = a(b-b)$. Отсюда, пользуясь дистрибутивностью умножения относительно вычитания, выводим $a0 = ab - ab$. Т. к. $ab - ab = 0$, получим $a0 = 0$ и в силу коммутативности умножения $0a = 0$.

7. Если e_1 и e_2 — единичные элементы, то $e_1e_2 = e_1$, т.к. e_2 — единичный элемент и $e_1e_2 = e_2$ т.к. e_1 — единичный элемент. Отсюда получим $e_1 = e_2$.

8. Если b_1 и b_2 — обратные элементы к a, то $b_1a = 1$, $ab_2 = 1$. Из ассоциативности умножения следует, что $(b_1a)b_2 = b_1(ab_2)$. Отсюда $1b_2 = b_11$ и $b_1 = b_2$.

10. Если $a = 0$, то утверждение верно. При $a \neq 0$ существует обратный a^{-1}. Умножим обе части равенства $ab = 0$ на a^{-1}. Получим $a^{-1}ab = a^{-1}0 \Rightarrow 1b = 0 \Rightarrow b = 0$. Следовательно, утверждение верно.

11. Противоположный a элемент $(-a)$ единственен. Сложим $a = 1a$ и $(-1)a$, получим $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$. Отсюда $(-1)a$ является противоположным a, т. е. $(-1)a = (-a)$.

12. $ab + (-a)b = (a + (-a))b = 0b = 0 \Rightarrow (-a)b = -(ab)$.

$ab + a(-b) = a(b + (-b)) = a0 = 0 \Rightarrow a(-b) = -(ab)$.

13. Если $c + (-c) = 0$, то $-(-c) = c$. Используя это и свойство 12, получим $ab = -(-(ab)) = -((-a)b) = (-a)(-b)$.

14. Доказывается как свойство 4.

15 а). Из свойства 14 имеем $\dfrac{a}{b} = b^{-1}a$, $\dfrac{c}{d} = cd^{-1}$. Тогда $\dfrac{a}{b} = \dfrac{c}{d} \Leftrightarrow b^{-1}a = cd^{-1} \Leftrightarrow b(b^{-1}a)d = b(cd^{-1})d \Leftrightarrow (bb^{-1})(ad) = (bc)(d^{-1}d) \Leftrightarrow 1(ad) = (bc)1 \Leftrightarrow ad = bc$.

15 б). Из свойства 15 а) $\dfrac{a}{b} = \dfrac{ac}{bc} \Leftrightarrow abc = bac$. Равенство $abc = bac$ верно, поэтому равенство $\dfrac{a}{b} = \dfrac{ac}{bc}$ также верно.

15 в). Из свойства 14 имеем $\dfrac{a}{b} = ab^{-1}$, $\dfrac{c}{d} = cd^{-1}$. Тогда $\dfrac{a}{b}\dfrac{c}{d} = ab^{-1}cd^{-1} = acb^{-1}d^{-1} = ac(db)^{-1} = ac(bd)^{-1} = \dfrac{ac}{bd}$.

15 г). Из свойства 14 имеем $\frac{a}{b} = ab^{-1}$, $\frac{c}{b} = cb^{-1}$. Тогда $\frac{a}{b} \pm \frac{c}{b} = ab^{-1} \pm cb^{-1} = (a \pm c)b^{-1} = \frac{a \pm c}{b}$.

15 д). Из свойства 15 б) имеем $\frac{a}{b} = \frac{ad}{bd}$, $\frac{c}{d} = \frac{cb}{bd}$. Далее из свойства 15 г) получаем $\frac{a}{b} \pm \frac{c}{d} = \frac{ad}{bd} \pm \frac{cb}{bd} = \frac{ad \pm cb}{bd}$. ■

Определение 2. *Множество M поля P называется* **подполем** *P, если оно само является полем при тех же операциях сложения и умножения, которые заданы в поле P. Тогда P называется надполем или расширением поля M.*

Теорема 1.8.2 (Критерий подполя). *Пусть заданы поле P и его подмножество M, содержащее не менее двух элементов.*

M является подполем поля P тогда и только тогда, когда P замкнуто относительно операций сложения, умножения, перехода к противоположному элементу и перехода к обратному элементу, т. е.

(КПП1) *для любых элементов* $a, b \in M$ $a + b \in M$;

(КПП2) *для любых элементов* $a, b \in M$ $ab \in M$;

(КПП3) *для любого элемента* $a \in M$ $(-a) \in M$;

(КПП4) *для любого элемента* $a \in M$, $a \neq 0$ $a^{-1} \in M$.

Доказательство.

1. Если M — подполе P, то M — поле и выполнены (КПП1) – (КПП4).

2. Пусть в M выполнены (КПП1) – (КПП4). Тогда в M определены операции сложения, умножения, любой элемент $a \in M$ имеет противоположный элемент $(-a) \in M$, любой элемент $a \in M$, $a \neq 0$ имеет обратный элемент $a^{-1} \in M$. Сложив элементы $a+(-a)$, получим $0 \in M$. Перемножив элементы aa^{-1}, получим $1 \in M$. Тогда аксиомы поля, выполненные в P, будут выполнены в M. Следовательно M — поле. ■

Задача. С помощью критерия подполя проверьте, что множество действительных чисел вида $a + b\sqrt{2}$ с любыми $a, b \in Q$ является полем.

Определение 3. *Характеристика поля — это наименьшее натуральное число n такое, что сумма n единиц равна нулю:*

$\sum\limits_{m=1}^{n} 1 = \underbrace{1+1+\ldots+1}_{n} = 0$. *Если такого числа не существует, то характеристика считается равной нулю.* ■

Теорема 1.8.3. *Характеристикой поля может быть либо 0, либо простое число.*

Доказательство. Если характеристика поля 0, то утверждение верно.

Пусть $n \in N$ — характеристика поля. Тогда n — это наименьшее натуральное число такое, что $\underbrace{1+1+\cdots+1}_{n} = 0$. Докажем, что n — простое число.

$\sqcap$ Пусть n — составное число. Тогда $n = pq$ $(p, q \in N,\ p < n,\ q < n)$. Из дистрибутивности $(\underbrace{1+1+\ldots+1}_{p})(\underbrace{1+1+\ldots+1}_{q}) = \underbrace{1\cdot 1+1\cdot 1+\ldots+1\cdot 1}_{pq}$ $= \underbrace{1+1+\ldots+1}_{n}$, т. е. $(\underbrace{1+1+\ldots+1}_{p})(\underbrace{1+1+\ldots+1}_{q}) = 0$. Отсюда следует, что $\underbrace{1+1+\ldots+1}_{p} = 0$, где $p < n$ или $\underbrace{1+1+\ldots+1}_{q} = 0$, где $q < n$. То есть n не является наименьшим натуральным числом, для которого $\underbrace{1+1+\ldots+1}_{n} = 0$. $\sqcup$ ■

Задачи.

Задача 1.8.1. *Докажите, что пересечение подполей будет подполем.*

Задача 1.8.2. *Докажите остальные свойства поля из теоремы 1.8.1.*

Рассмотрим множества $\mathbb{N}$ – натуральных чисел, $\mathbb{Z}$ – целых чисел, $\mathbb{Q}$ – рациональных чисел, $\mathbb{R}$ – действительных чисел, $\mathbb{C}$ – комплексных чисел.

Далее везде считаем, что для $\mathbb{R}$ выполнены условия A1–A9, указанные в определения 1, и, следовательно, в замечаниях 1–4, а также в теореме 1.8.1. Следовательно, $\mathbb{R}$ с операциями сложения и умножения — поле.

Задача 1.8.3.

а) Докажите, что $\mathbb{N}$ *и* $\mathbb{Z}$ *с операциями сложения и умножения не являются полями.*

б) Докажите, что $\mathbb{Q}$ *с операциями сложения и умножения является полем.*

в) Докажите, что $\mathbb{C}$ *с операциями сложения и умножения является полем.*

Задача 1.8.4. *Проверьте является ли кольцом или полем данное множество с указанными операциями.*

а) Матрицы $\mathbb{R}^{m\times n}$ *(размера* $m \times n$ *из вещественных элементов) с операциями сложения и умножения.*

б) Матрицы $\mathbb{R}^{n\times n}$ *с операциями сложения и умножения.*

в) Невырожденные матрицы из $\mathbb{R}^{n\times n}$ *(определители матриц не равны нулю) с операциями сложения и умножения.*

г) Многочлены $\mathbb{R}[x]$ *(с действительными коэффициентами от действительной переменной* x *) с операциями сложения и умножения.*

д) Рациональные функции или дроби вида $\dfrac{f(x)}{g(x)}$, *где* $f(x), g(x) \in \mathbb{R}[x]$, $g(x) \not\equiv 0$, *с операциями сложения* $\dfrac{f(x)}{g(x)}+\dfrac{h(x)}{k(x)} = \dfrac{f(x)k(x)+g(x)h(x)}{g(x)k(x)}$ *и умножения* $\dfrac{f(x)}{g(x)} \cdot \dfrac{h(x)}{k(x)} = \dfrac{f(x)h(x)}{g(x)k(x)}$.

1.9. Кольцо классов вычетов

Z образует абелеву группу по сложению, mZ — это подгруппа группы Z. Подгруппа mZ является нормальным делителем, т. к. совпадают левые и правые смежные классы, порождённые любым элементом $a \in Z:\ a + mZ = mZ + a$. Смежный класс $a + mZ = \{a + mk \mid k \in Z\}$, содержащий число a, называется *классом вычетов по модулю* m и обозначается $[a]_m$. Все числа из класса $[a]_m$ сравнимы по модулю m, т. к. они при делении на m дают такой же остаток как и a. Поэтому всего имеется m различных классов вычетов $Z_m = \{[0]_m, [1]_m, \ldots, [m-1]_m\}$. Они образуют (см. теорему 1.6.2) группу с операцией сложения классов

$$[a]_m + [b]_m = [a+b]_m. \tag{1}$$

Введём операцию умножения классов

$$[a]_m[b]_m = [ab]_m. \tag{2}$$

Теорема 1.9.1. *Z_m (классы вычетов по модулю m) с операциями (1) и (2) являются коммутативным кольцом с единицей.*

Доказательство.

АК1 Сложение ассоциативно $([a]_m+[b]_m)+[c]_m = [a]_m+([b]_m+[c]_m)$,

т. к. $([a]_m + [b]_m) + [c]_m = [a+b]_m + [c]_m = [(a+b)+c]_m = [a+(b+c)]_m$,

$[a]_m + ([b]_m + [c]_m) = [a]_m + [b+c]_m = [a+(b+c)]_m$.

АК4 Сложение коммутативно $[a+b]_m = [b+a]_m$, т. к. $a+b = b+a$.

АК2 $[0]_m$ — нейтральный элемент для сложения,

т. к. $[a]_m + [0]_m = [a+0]_m = [a]_m$.

АК3 $[-a]_m$ — противоположный элемент для $[a]_m$,

т. к. $[a]_m + [-a]_m = [a+(-a)]_m = [0]_m$.

АК8 Умножение коммутативно $[a]_m[b]_m = [b]_m[a]_m$,

т. к. $[a]_m[b]_m = [ab]_m = [ba]_m = [ab]_m$.

АК5 Умножение ассоциативно $([a]_m[b]_m)[c]_m = [a]_m([b]_m[c]_m)$,

т. к. $([a]_m[b]_m)[c]_m = [ab]_m[c]_m = [(ab)c]_m = [a(bc)]_m = [a]_m[bc]_m = [a]_m([b]_m[c]_m)$. ▮

АК6 Умножение дистрибутивно $([a]_m+[b]_m)[c]_m = [a]_m[c]_m+[b]_m[c]_m$, ▮

т. к. $([a]_m+[b]_m)[c]_m = [a+b]_m[c]_m = [(a+b)c]_m = [ac+bc]_m = [ac]_m+[bc]_m =$
$= [a]_m[c]_m + [b]_m[c]_m$.

АК7 $[1]_m$ — нейтральный элемент для умножения,

т. к. $[a]_m[1]_m = [a1]_m = [a]_m$. ■

Приведём три примера конечных колец, указывая отдельно таблицы сложения и умножения и обозначая классы вычетов их представителями.

$\mathbf{Z_2}$:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

$\mathbf{Z_3}$:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

+	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

$\mathbf{Z_4}$:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

+	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

В кольце $\mathbf{Z_4}$ есть делители нуля: $[2]_4[2]_4 = [0]_4$ — «$2 \times 2{=}0$».

Теорема 1.9.2. *Кольцо Z_m является полем тогда и только тогда, когда m — простое число. При этом для простого числа m поле Z_m имеет характеристику m.*

Доказательство. I. Пусть кольцо Z_m является полем. Если $m = 2$, то это простое число. При $m > 2$ предположим, что m — составное число, т. е. $m = pq$ ($p, q \in N, 1 < p < m, 1 < q < m$). Тогда $[m]_m = [pq]_m = [p]_m[q]_m$, где $[m]_m = [0]_m$ — нулевой класс. Отсюда $[p]_m[q]_m = [0]_m$ для $[p]_m \neq [0]_m$, $[q]_m \neq [0]_m$. Т. е. в поле Z_m есть делители нуля, что невозможно. Следовательно, m — простое число.

II. Пусть m — простое число. В теореме 1.9.1 показано, что Z_m — коммутативное кольцо с единицей. Для того чтобы доказать, что Z_m — поле, осталось проверить наличие у произвольного ненулевого класса $[a]_m (1 \leqslant a \leqslant m-1)$ обратного класса $[b]_m$: $[b]_m[a]_m = [1]_m$. Для этого рассмотрим m классов $[0]_m[a]_m = [0a]_m$, $[1]_m[a]_m = [1a]_m, \ldots, [m-1]_m[a]_m = [(m-1)a]_m$ и сначала покажем, что все они различны. Действительно, из совпадения классов $[pa]_m = [qa]_m$, где p и q различны, $1 \leqslant a \leqslant m-1$, $0 \leqslant p \leqslant m-1$, $0 \leqslant q \leqslant m-1$, следует, что 1) $pa - qa = (p-q)a$ делится на m, 2) $1 \leqslant a \leqslant m-1$, 3) $-(m-1) \leqslant p-q \leqslant (m-1)$. Число a не делится на m в силу п. 2). Тогда из простоты m и п. 1) следует, что $p-q$ делится на m. При выполнении п. 3) $p-q$ делится на m только при $p-q=0$, т.е. при $p=q$, что противоречит условию.

Поскольку все m классов $[0a]_m, [1a]_m, \ldots, [(m-1)a]_m$ различны, то среди них найдётся класс $[ba]_m$ равный $[1]_m$, и тогда $[b]_m[a]_m = [ba]_m = [1]_m$. Поэтому при простом m у произвольного ненулевого класса $[a]_m$ есть обратный $[b]_m$, и, следовательно, Z_m — поле. Это поле характе-

ристики m, т. к. $\underbrace{[1]_m + \ldots + [1]_m}_{p} = [p]_m \neq [0]_m$ при $0 < p < m$ и $\underbrace{[1]_m + \ldots + [1]_m}_{m} = [m]_m = [0]_m$. ■

1.10. Изоморфизм групп, колец, полей

Определение 1. *Группа* $\langle \Gamma_1, \cdot \rangle$ *изоморфна группе* $\langle \Gamma_2, \odot \rangle$, *если существует биективное отображение* $f\colon \Gamma_1 \to \Gamma_2$ *такое, что*

$$\forall a, b \in \Gamma_1 \quad f(a \cdot b) = f(a) \odot f(b),$$

т. е. образ произведения $f(a \cdot b)$ *равен произведению образов* $f(a) \odot f(b)$.

Обозначение: $\Gamma_1 \cong \Gamma_2$. *При этом говорят, что* f *— изоморфизм.* ■

Примеры.

1. Группа целых чисел $\langle Z, + \rangle$ изоморфна группе чётных чисел $\langle 2Z, + \rangle$. Отображение $f\colon Z \to 2Z$, $f(n) = 2n$, $n \in Z$ есть изоморфизм: 1) f — биекция, 2) $\forall n_1, n_2 \in Z \quad f(n_1 + n_2) = 2(n_1 + n_2) = 2n_1 + 2n_2 = = f(n_1) + f(n_2)$.

2. Пусть $\Gamma = \{x | x \in R,\ x > 0\}$, тогда $\langle \Gamma, \cdot \rangle \cong \langle R, + \rangle$. Изоморфизм $f\colon \Gamma \to R$, $f(x) = \ln(x)$, т.к. 1) f — биекция, 2) $f(x_1 x_2) = \ln(x_1 x_2) = \ln(x_1) + \ln(x_2) = f(x_1) + f(x_2)$. ■

Теорема 1.10.1 (Свойства изоморфизма групп).

Если группа $\langle \Gamma_1, \cdot \rangle$ *изоморфна группе* $\langle \Gamma_2, \odot \rangle$ *и* $f\colon \Gamma_1 \to \Gamma_2$ *— изоморфизм, то выполнены следующие свойства.*

1° $f^{-1}\colon \Gamma_2 \to \Gamma_1$ *— также изоморфизм.*

2° *Образ единицы* $e_1 \in \Gamma_1$ *равен единице* $e_2 \in \Gamma_2$: $\quad f(e_1) = e_2$.

3° *Образ обратного элемента для* $a \in \Gamma_1$ *является обратным для образа* $f(a)$, *т. е.* $f(a^{-1}) = (f(a))^{-1}$.

Доказательство.

1° Если $f\colon \Gamma_1 \to \Gamma_2$ — биекция, то существует обратное отображение $f^{-1}\colon \Gamma_2 \to \Gamma_1$ также биективное. Покажем, что

$$\forall c, d \in \Gamma_2 \quad f^{-1}(c \odot d) = f^{-1}(c) \cdot f^{-1}(d).$$

Обозначим $a = f^{-1}(c)$, $b = f^{-1}(d)$. Тогда $a, b \in \Gamma_1$ и $f(a) = c$, $f(b) = d$. Поскольку f — изоморфизм, то $f(a \cdot b) = f(a) \odot f(b) = c \odot d$, т. е. $c \odot d = f(a \cdot b)$. Применим f^{-1} и получим $f^{-1}(c \odot d) = f^{-1}(f(a \cdot b))$. Отсюда $f^{-1}(c \odot d) = a \cdot b$, т. е. $f^{-1}(c \odot d) = f^{-1}(c) \cdot f^{-1}(d)$. Следовательно, $f^{-1} \colon \Gamma_2 \to \Gamma_1$ — изоморфизм.

2° Применим к обеим частям равенства $e_1 = e_1 \cdot e_1$ изоморфизм f, получим $f(e_1) = f(e_1 \cdot e_1)$ и $f(e_1) = f(e_1) \odot f(e_1)$. Представим левую часть в виде $f(e_1) \odot e_2$, тогда $f(e_1) \odot e_2 = f(e_1) \odot f(e_1)$. Сократив на $f(e_1)$, получим $e_2 = f(e_1)$.

3° Применим к обеим частям равенства $a \cdot a^{-1} = e_1$ изоморфизм f, и получим $f(a \cdot a^{-1}) = f(e_1) \Rightarrow f(a) \odot f(a^{-1}) = e_2$. Отсюда $f(a^{-1}) = (f(a))^{-1}$. ■

Определение 2. *Кольцо* $\langle K_1, +, \cdot, \rangle$ *изоморфно кольцу* $\langle K_2, \oplus, \odot \rangle$, *если существует биективное отображение* $f \colon K_1 \to K_2$ *такое, что*

$$\forall a, b \in K_1 \quad f(a \cdot b) = f(a) \odot f(b) \ \text{ и } \ f(a + b) = f(a) \oplus f(b).$$

Обозначение: $K_1 \cong K_2$. *При этом говорят, что* f — *изоморфизм.* ■

Теорема 1.10.2 (Свойства изоморфизма колец).

Если кольцо $\langle K_1, +, \cdot \rangle$ *изоморфно кольцу* $\langle K_2, \oplus, \odot \rangle$ *и* $f \colon K_1 \to K_2$ — *изоморфизм, то выполнены следующие свойства.*

1° $f^{-1} \colon K_2 \to K_1$ — *также изоморфизм.*

2° *Образ нуля* $0_1 \in K_1$ *равен нулю* $0_2 \in K_2$: $f(0_1) = 0_2$.

3° *Образ противоположного элемента для* $a \in K_1$ *является противоположным для образа* $f(a) \in K_2$, *т. е.* $f(-a) = -f(a)$.

Доказательство проведите самостоятельно. ■

Определение 3. *Поле* $\langle P_1, +, \cdot, \rangle$ *изоморфно полю* $\langle P_2, \oplus, \odot, \rangle$, *если существует биективное отображение* $f \colon P_1 \to P_2$ *такое, что*

$$\forall a, b \in P_1 \quad f(a \cdot b) = f(a) \odot f(b) \ \text{ и } \ f(a + b) = f(a) \oplus f(b).$$

Обозначение: $P_1 \cong P_2$. *При этом говорят, что* f — *изоморфизм.* ■

Теорема 1.10.3 (Свойства изоморфизма полей).

Если поле $\langle P_1, +, \cdot \rangle$ *изоморфно полю* $\langle P_2, \oplus, \odot \rangle$ *и* $f\colon P_1 \to P_2$ — *изоморфизм, то выполнены следующие свойства.*

1° $f^{-1}\colon P_2 \to P_1$ — *также изоморфизм.*

2° *Образ нуля* $0_1 \in P_1$ *равен нулю* $0_2 \in P_2$: $f(0_1) = 0_2$.

3° *Образ противоположного элемента для* $a \in P_1$ *является противоположным для образа* $f(a) \in P_2$, *т. е.* $f(-a) = -f(a)$.

4° *Образ единицы* $1_1 \in P_1$ *равен единице* $1_2 \in P_2$: $f(1_1) = 1_2$.

5° *Образ обратного элемента для* $a \in P_1$, $a \neq 0_1$ *является обратным для образа* $f(a) \in P_2$, *т. е.* $f(a^{-1}) = (f(a))^{-1}$. ■

Доказательство проведите самостоятельно. ■

2. Комплексные числа

2.1. Комплексные числа. Поле комплексных чисел

Определение 1.

Комплексными числами называются упорядоченные пары (x, y) *вещественных чисел* $x, y \in R$. *Обозначение:* $z = (x, y)$, $z \in C$, *где* C — *множество всех комплексных чисел.*

1) *Два комплексных числа* $z_1 = (x_1, y_1)$ *и* $z_2 = (x_2, y_2)$ *равны, если* $x_1 = x_2$, $y_1 = y_2$; (*обозначение:* $z_1 = z_2$).

2) *Суммой чисел* $z_1 = (x_1, y_1) \in C$ *и* $z_2 = (x_2, y_2) \in C$ *называется число* $(x_1 + x_2, y_1 + y_2) \in C$ (*обозначение:* $z_1 + z_2$). *Операция* $\langle z_1, z_2 \rangle \mapsto z_1 + z_2$ *называется сложением.*

3) *Произведением чисел* $z_1 = (x_1, y_1) \in C$ *и* $z_2 = (x_2, y_2) \in C$ *называется число* $(x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \in C$ (*обозначение:* $z_1 \cdot z_2$ *или* z_1z_2). *Операция* $\langle z_1, z_2 \rangle \mapsto z_1z_2$ *называется умножением.*

4) *Обозначим через* C_R *подмножество комплексных чисел вида* $(x, 0)$. *Каждое число* $(x, 0) \in C_R$ *взаимно-однозначно соответствуют вещественному числу* $x \in R$. *Поэтому числа из* C_R *называют вещественными числами.* ■

Теорема 2.1.1. *Множество комплексных чисел C с операциями сложения и умножения являются полем.*

Доказательство.

Из определения следует, что операции сложения и умножения корректны.

Для $z_k = (x_k, y_k) \in C$, $k = \overline{1,2}$ из $z_1+z_2 = (x_1+x_2, y_1+y_2)$, $z_2+z_1 = (x_2+x_1, y_2+y_1)$, $z_1z_2 = (x_1x_2-y_1y_2, x_1y_2+y_1x_2)$, $z_2z_1 = (x_2x_1-y_2y_1, x_2y_1+ +y_2x_1)$ следует $z_1+z_2 = z_2+z_1$, $z_1z_2 = z_2z_1$ — коммутативность сложения и умножения АП4 и АП8.

Для $z_k = (x_k, y_k) \in C$, $k = \overline{1,3}$ из $(z_1 + z_2) + z_3 = (x_1 + x_2, y_1 + + y_2) + (x_3, y_3) = ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3)$, $z_1 + (z_2 + z_3) = (x_1, y_1) + +(x_2+x_3, y_2+y_3) = (x_1+(x_2+x_3), y_1+(y_2+y_3))$ следует ассоциативность сложения АП1.

АП2: нулём является число $(0,0)$, т. к. $\forall\,(x,y) \in C$ $(x,y) + (0,0) = = (x,y)$.

АП3: противоположным числу $(x,y) \in C$ является число $(-x,-y)$, т. к. $(x,y) + (-x,-y) = (0,0)$.

АП5: ассоциативность умножения $(z_1z_2)z_3 = z_1(z_2z_3)$ проверьте самостоятельно.

АП6. Найдём $w = (u,v) \in C$, для которого $\forall\, z = (x,y) \in C$ $zw = = z \Leftrightarrow (x,y)(u,v) = (x,y) \Leftrightarrow (xu-yv, xv+yu) = (x,y) \Leftrightarrow \begin{cases} xu - yv = x, \\ yu + xv = y. \end{cases}$

Определитель системы $\Delta = \begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2$. При $x \neq 0$ или $y \neq 0$ система имеет единственное решение. Находим определители $\Delta_1 = \begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2$, $\Delta_2 = \begin{vmatrix} x & x \\ y & y \end{vmatrix} = 0$ и решение $u = \dfrac{\Delta_1}{\Delta} = 1$, $v = \dfrac{\Delta_2}{\Delta} = 0$. При $x = 0$ и $y = 0$ решение любое. Итак, $w = (1,0)$.

АП7: найдём для $(x,y) \neq (0,0)$ такое (u,v), что $(x,y)(u,v) = = (1,0) \Leftrightarrow (xu-yv, xv+yu) = (1,0) \Leftrightarrow \begin{cases} xu - yv = 1, \\ xv + yu = 0, \end{cases} \Leftrightarrow \begin{cases} xu - yv = 1, \\ yu + xv = 0. \end{cases}$

Определитель системы $\Delta = \begin{vmatrix} x & -y \\ y & x \end{vmatrix} = x^2 + y^2 \neq 0$. Следователь-

но, система имеет единственное решение. Находим определители $\Delta_1 = \begin{vmatrix} 1 & -y \\ 0 & x \end{vmatrix} = x$, $\Delta_2 = \begin{vmatrix} x & 1 \\ y & 0 \end{vmatrix} = -y$ и решение $u = \dfrac{\Delta_1}{\Delta} = \dfrac{x}{x^2+y^2}$, $v = \dfrac{\Delta_2}{\Delta} = \dfrac{-y}{x^2+y^2}$.

Итак, $(x,y)^{-1} = \left(\dfrac{x}{x^2+y^2}, \dfrac{-y}{x^2+y^2}\right)$.

АП9: дистрибутивность умножения относительно сложения $(z_1 + z_2)z_3 = z_1z_3 + z_2z_3$ проверьте самостоятельно.

Следствия. 1. Для любой пары комплексных чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$ существует единственная разность $z_1 - z_2 = (x_1 - x_2, y_1 - y_2)$.

2. Для любой пары комплексных чисел $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2) \neq$ $\neq (0,0)$ существует единственное частное

$$\frac{z_1}{z_2} = z_1{z_2}^{-1} = \left(\frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2}, \frac{-x_1y_2 + y_1x_2}{x_2^2 + y_2^2}\right).$$

3. Подмножество комплексных чисел $C_R = \{z \,|\, z = (x, 0),\ x \in R\}$ является подполем поля C с операциями сложения и умножения комплексных чисел и это подполе изоморфно полю вещественных чисел R с операциями сложения и умножения вещественных чисел, причём изоморфизмом является отображение $f((x,0)) = x$, $x \in R$.

Доказательство. Следствие 1 получено из 4-го свойства поля. Для следствия 2 формула $\dfrac{z_1}{z_2} = z_1{z_2}^{-1}$ получена из 14-го свойства поля. Далее значение $z_1{z_2}^{-1} = \left(\dfrac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2}, \dfrac{-x_1y_2 + y_1x_2}{x_2^2 + y_2^2}\right)$ получено при умножении $z_1 = (x_1, y_1)$ на обратное число $z_2^{-1} = \left(\dfrac{x_2}{x_2^2 + y_2^2}, \dfrac{-y_2}{x_2^2 + y_2^2}\right)$.

3. Действительно, для C_R выполнены критерии подполя.

(КП1) $\forall\, z_1 = (x_1, 0),\ z_2 = (x_2, 0) \in C_R \quad z_1 + z_2 = (x_1 + x_2, 0) \in C_R$;

(КП2) $\forall\, z_1 = (x_1, 0),\ z_2 = (x_2, 0) \in C_R \ \ z_1z_2 = (x_1x_2 - 0 \cdot 0, x_10 + 0x_2) =$ $= (x_1x_2, 0) \in C_R$;

(КП3) $\forall\, z = (x, 0) \in C_R \quad -z = (-x, 0) \in C_R$;

(КП4) $\forall\, z = (x, 0) \in C_R,\ x \neq 0 \ \ z^{-1} = \left(\dfrac{x}{x^2 + 0^2}, \dfrac{-0}{x^2 + 0^2}\right) = \left(\dfrac{1}{x}, 0\right)$,

причём $\left(\frac{1}{x}, 0\right) \in C_R$.

Отображение $f\colon C_R \to R$, действующее по правилу $f((x,0)) = x$, является изоморфизмом, так как а) f биективно, б) $\forall\, z_1 = (x_1, 0)$, $z_2 = (x_2, 0) \in C_R$ $f(z_1 + z_2) = f((x_1 + x_2, 0)) = x_1 + x_2 = f(z_1) + f(z_2)$ и $f(z_1 z_2) = f((x_1 x_2, 0)) = x_1 x_2 = f(z_1) f(z_2)$. ■

2.2. Алгебраическая форма комплексных чисел. Арифметические операции для комплексных чисел в алгебраической форме. Геометрическое изображение комплексных чисел и операций сложения и вычитания

Определение 1. *Числа вида* $(x, 0) \in C$ *называются* **вещественными** *и обозначаются просто* x.

Число $(0, 1) \in C$ *называется* **мнимая единица** *и обозначается* i. ■

Отсюда и из определения комплексных чисел получаем следующие факты.

1° $i^2 = -1$.

2° Любое комплексное число $z = (x, y)$ может быть записано в виде

$$z = x + yi.$$

Доказательство.

1° $i^2 = ii = (0,1)\,(0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1$.

2° Имеем $z = (x, y) = (x, 0) + (0, y)$. Далее получим $(y, 0)\,(0, 1) = (y \cdot 0 - 0 \cdot 1, y \cdot 1 + 0 \cdot 0) = (0, y)$. Отсюда $(x, 0) + (0, y) = (x, 0) + (y, 0)\,(0, 1) = x + yi$.

Определение 2. *Форма записи комплексного числа* $z = (x, y)$ *в виде* $z = x + yi$ *называется* **алгебраической формой** *комплексного числа. При этом* x *и* y *называются соответственно* **вещественной и мнимой** *частью числа* z *и обозначаются* $x = \mathrm{Re}(z)$, $y = \mathrm{Im}(z)$. *Отметим, что* $\mathrm{Re}(z) \in R$, $\mathrm{Im}(z) \in R$. *Числа вида* $z = x + 0i$ *называют*

вещественными и записывают в виде $z = x$. *Числа вида* $z = 0 + yi$ *называют чисто мнимыми и записывают в виде* $z = yi$.

Два комплексных числа $z_1 = x_1 + y_1 i$ *и* $z_2 = x_2 + y_2 i$ *равны тогда и только тогда, когда равны их действительные и мнимые части:* $x_1 = x_2$ *и* $y_1 = y_2$.

Комплексное число $\overline{z} = x - yi$ *называется сопряжённым к числу* $z = x + yi$.

Преобразование $z \mapsto \overline{z}$ *называется операцией сопряжения.* ■

Арифметические операции для комплексных чисел в алгебраической форме.

Пусть заданы два комплексных числа $z_1 = x_1 + y_1 i$ и $z_2 = x_2 + y_2 i$. При выполнении алгебраических операций используем ассоциативность, коммутативность и дистрибутивность сложения и умножения.

Тогда операции сложения и вычитания можно производить почленно отдельно для вещественных и мнимых частей:

$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i,$

$z_1 - z_2 = (x_1 - x_2) + (y_1 - y_2)i.$

Операцию умножения можно произвести так:

$z_1 z_2 = (x_1 + y_1 i)(x_2 + y_2 i) = x_1 x_2 + x_1 y_2 i + y_1 x_2 i + y_1 y_2 i^2$, учитывая равенство $i^2 = -1$, далее получим $(x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i$.

Для деления $\dfrac{z_1}{z_2}$ удобно числитель и знаменатель дроби умножить на сопряжённое к знаменателю число $\overline{z_2} = x_2 - y_2 i$:

$$\frac{z_1}{z_2} = \frac{x_1 + y_1 i}{x_2 + y_2 i} = \frac{(x_1 + y_1 i)(x_2 - y_2 i)}{(x_2 + y_2 i)(x_2 - y_2 i)} = \frac{(x_1 x_2 + y_1 y_2) + (y_1 x_2 - x_1 y_2)i}{x_2^2 + y_2^2} =$$

$$= \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + \frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2} i. \blacksquare$$

Примеры.

1) $(2 + 4i) + (-3 - 5i) = (2 - 3) + (4 - 5)i = -1 - i;$

2) $(-7 + 3i) - (-3 + 5i) = (-7 + 3) + (3 - 5)i = -4 - 2i;$

3) $(-2 - 3i)(3 + 2i) = -6 - 4i - 9i - 6i^2 = -6 - 13i + 6 = -13i.$

4) $(5 - 4i)(-3 - 2i) = -15 - 8 + (-10 + 12)i = -23 + 2i.$

5) $\dfrac{3-7i}{-2+3i} = \dfrac{(3-7i)(-2-3i)}{(-2+3i)(-2-3i)} = \dfrac{-6-21+(-9+14)i}{4+9} = \dfrac{-27+5i}{13} =$
$= \dfrac{-27}{13} + \dfrac{5}{13}i;$

6) $\dfrac{5-10i}{4-3i} = \dfrac{5(1-2i)}{4-3i} = \dfrac{5(1-2i)(4+3i)}{(4-3i)(4+3i)} = \dfrac{5(4+6+(3-8)i)}{16+9} =$
$= \dfrac{5(10-5i)}{25} = \dfrac{25(2-i)}{25} = 2-i.$

Геометрическое изображение комплексных чисел и операций сложения и вычитания.

Комплексное число $z = x + yi$ изображается точкой M или вектором $\overline{OM}$ плоскости с декартовыми прямоугольными координатами (x, y) (рис. 7 для случая $x > 0$ и $y > 0$). Тогда сопряжённое число $\overline{z} = x - yi$ изображается точкой или вектором с декартовыми прямоугольными координатами $(x, -y)$, противоположное число $-z = -x - yi$ изображается точкой или вектором с декартовыми прямоугольными координатами $(-x, -y)$.

Вещественные числа изображаются точками оси абсцисс, чисто мнимые числа изображаются точками оси ординат. Число ноль изображается началом координат. Сопряжённое число для данного числа изображается точкой, симметричной данной точке относительно оси абсцисс.

Плоскость, точками которой изображаются комплексные числа называется *комплексной плоскостью*, её ось абсцисс называется *вещественной осью*, ось ординат — *мнимой осью*.

Сумму $z_1 + z_2$ и разность $z_1 - z_2 = z_1 + (-z_2)$ удобно выражать радиус-векторами, получаемыми по правилу параллелограмма или по правилу треугольника (рис. 8).

2.3. Тригонометрическая форма комплексного числа. Операции над комплексными числами в тригонометрической форме

Комплексное число $z = x + yi$ изображается точкой M или вектором $\overline{OM}$ плоскости с декартовыми прямоугольными координатами (x, y).

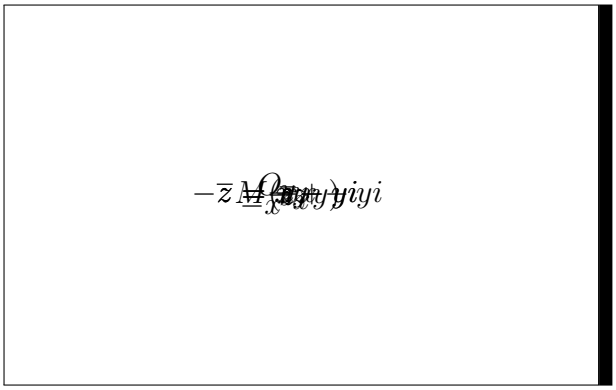

Рис. 7

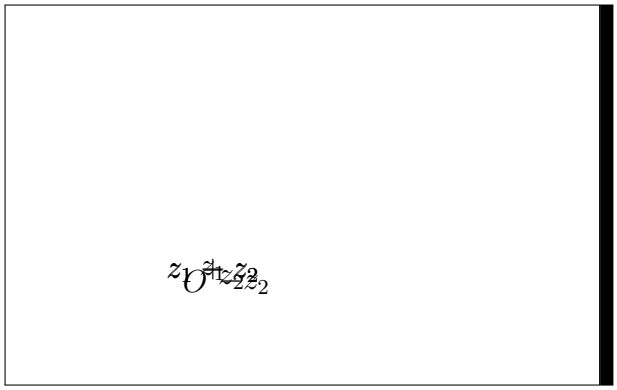

Рис. 8

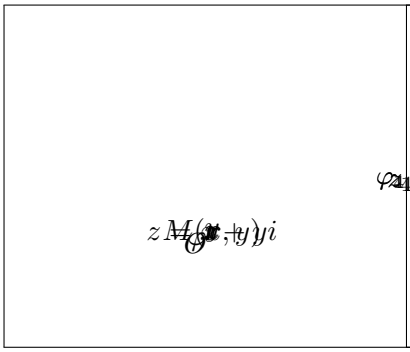

Рис. 9

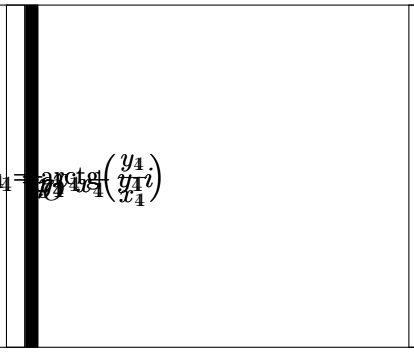

Рис. 10

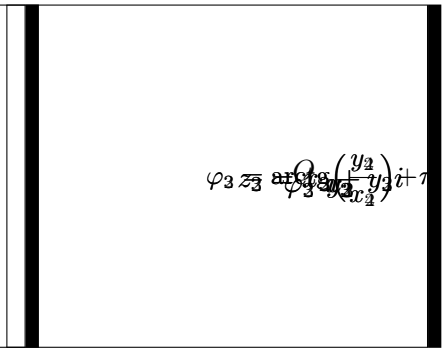

Рис. 11

Положение точки M на плоскости вполне определяется также её полярными координатами: длиной r радиус-вектора $\overline{OM}$ и углом φ поворота от положительной полуоси абсцисс до вектора $\overline{OM}$ (рис. 9). Неотрицательное вещественное число $r = \sqrt{x^2 + y^2}$ называется **модулем** числа $z = x + yi$ и обозначается $|z|$. Итак, $|x + yi| = \sqrt{x^2 + y^2}$. Модуль равен нулю только для числа $z = 0$. Угол φ называется **аргументом** числа $z = x + yi$ и обозначается $\arg(z)$. Положительные углы отсчитываются против часовой стрелки, отрицательные — по часовой стрелке. Аргумент любого числа $z \neq 0$ определён с точностью до числа кратного 2π : $\arg(z) = \varphi_0 + 2\pi k$, $k \in Z$. Аргумент нуля не определён, но удобно считать, что аргументом нуля является произвольное число.

Между декартовыми прямоугольными координатами (x, y) и полярными координатами (r, φ) точки существует следующая связь: $\begin{cases} x = r\cos\varphi, \\ y = r\sin\varphi. \end{cases}$

Применив эти формулы к произвольному комплексному числу $z = x + yi$, получим запись этого числа в виде

$$z = r(\cos\varphi + i\sin\varphi), \text{ где } r = |z|,\ \varphi = arg(z).$$

Эту запись называют **тригонометрической формой** комплексного числа.

Для нахождения r, φ по заданным x, y возведём x и y в квадрат, получим $x^2 + y^2 = r^2$ и, следовательно, $r = \sqrt{x^2+y^2}$. Если $x = 0$ и $y = 0$, то $r = 0$ и угол φ не определён. Если $x \neq 0$ или $y \neq 0$, то $r > 0$ и тогда угол φ определяется из равенств $\cos\varphi = \dfrac{x}{r} = \dfrac{x}{\sqrt{x^2+y^2}}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{y}{\sqrt{x^2+y^2}}$. Для получения φ в промежутке $(-\pi, \pi]$ поступим следующим образом. Возьмём $\varphi = \pi/2$ при $x = 0$, $y > 0$, $\varphi = -\pi/2$ при $x = 0$, $y < 0$. При $x \neq 0$ разделим y на x и получим $\operatorname{tg}\varphi = y/x$. Отсюда $\varphi = \operatorname{arctg}(y/x) + \pi k$, $k \in Z$. Далее берём $k = 0$, $\varphi = \operatorname{arctg}(y/x)$ при $x > 0$, т. е. для точек (x, y) из 1-й и 4-й четверти, $k = 1$, $\varphi = \operatorname{arctg}(y/x) + \pi$ при $x < 0$, $y \geqslant 0$, т. е. для точек (x, y) из 2-й четверти, $k = -1$, $\varphi = \operatorname{arctg}(y/x) - \pi$ при $x < 0$, $y < 0$, т. е. для точек (x, y) из 3-й четверти. См. рис. 10–11. ■

Теорема 2.3.1 (Операции в тригонометрической форме).

1° *Пусть* $z_k = r_k(\cos\varphi_k + i\sin\varphi_k) \in C$, $r_k = |z_k|$, $\varphi_k = \arg z_k$, $k = \overline{1,2}$, *тогда*

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)), \tag{3}$$

т.е. при умножении комплексных чисел их модули умножаются, а аргументы складываются: $|z_1 z_2| = |z_1||z_2|$, $\arg(z_1 z_2) = \arg(z_1) + arg(z_2)$.

2° *Пусть* $z = r(\cos\varphi + i\sin\varphi) \in C$, $r = |z|$, $\varphi = \arg z$, $z \neq 0$, *тогда*

$$z^{-1} = \frac{1}{z} = \frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)). \tag{4}$$

3° *Пусть* $z_k = r_k(\cos\varphi_k + i\sin\varphi_k) \in C$, $r_k = |z_k|$, $\varphi_k = \arg z_k$, $k = \overline{1,2}$, $z_2 \neq 0$, *тогда*

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)), \tag{5}$$

т.е. при делении комплексных чисел их модули делятся, а аргументы

вычитаются: $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$, $\arg\left(\dfrac{z_1}{z_2}\right) = \arg(z_1) - arg(z_2)$.

4° *Пусть* $z = r(\cos\varphi + i\sin\varphi) \in C$, $r = |z|$, $\varphi = \arg z$, *тогда*

$$(-z) = r(\cos(\varphi+\pi) + i\sin(\varphi+\pi)), \tag{6}$$

$$\overline{z} = r(\cos(-\varphi) + i\sin(-\varphi)). \tag{7}$$

Доказательство. 1° $z_1z_2 = r_1(\cos\varphi_1 + i\sin\varphi_1)r_2(\cos\varphi_2 + i\sin\varphi_2) = r_1r_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) = r_1r_2(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2) = r_1r_2(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2))$.

2° Используя формулу (3), найдём произведение чисел $r(\cos\varphi + i\sin\varphi)$ и $\dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)$, получим $\dfrac{r}{r}(\cos(\varphi-\varphi) + i\sin(\varphi-\varphi)) = \cos 0 + i\sin 0 = 1$. Следовательно, обратным для $r(\cos\varphi + i\sin\varphi)$ является $\dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)$.

3° В поле C выполнено равенство $\dfrac{z_1}{z_2} = z_1z_2^{-1}$. Используя формулы (3) и (4), получим $z_1z_2^{-1} = r_1(\cos\varphi_1 + i\sin\varphi_1)\dfrac{1}{r_2}(\cos(-\varphi_2) + i\sin(-\varphi_2)) = \dfrac{r_1}{r_2}(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2))$.

4° 1) $r(\cos(\varphi+\pi) + i\sin(\varphi+\pi)) = r(-\cos\varphi - i\sin\varphi) = -z$;

2) $r(\cos(-\varphi) + i\sin(-\varphi)) = r(\cos\varphi - i\sin\varphi) = \overline{z}$. ■

Следствие. *Пусть* $z_k = r_k(\cos\varphi_k + i\sin\varphi_k) \in C$, $r_k = |z_k|$, $\varphi_k = \arg z_k, k = \overline{1,n}$, *тогда*

$$z_1z_2\ldots z_n = (r_1r_2\ldots r_n)\big(\cos(\varphi_1+\varphi_2+\ldots+\varphi_n) + i\sin(\varphi_1+\varphi_2+\ldots+\varphi_n)\big), \tag{8}$$

т.е. при умножении чисел их модули перемножаются, а аргументы складываются.

Доказательство.

Докажем методом математической индукции.

При $n = 1$ формула $r_1(\cos\varphi_1 + i\sin\varphi_1) = r_1(\cos\varphi_1 + i\sin\varphi_1)$ верна.

Пусть при $n = m$ формула (8) верна.

Докажем формулу (8) при $n = m+1$:

$z_1\ldots z_mz_{m+1} = \big(z_1\ldots z_m\big)z_{m+1} = \big((r_1\ldots r_m)(\cos(\varphi_1+\ldots+\varphi_m) + i\sin(\varphi_1+\ldots+\varphi_m)\big)\big(r_{m+1}(\cos\varphi_{m+1} + i\sin\varphi_{m+1})\big) = \big(r_1\ldots r_mr_{m+1}\big)\big(\cos(\varphi_1+\ldots+$

$\varphi_m + \varphi_{m+1}) + i\sin(\varphi_1 + \ldots + \varphi_m + \varphi_{m+1})\big)$. Следовательно, формула (8) верна. ■

Из теоремы 2.3.1 получим геометрический смысл умножения и деления двух комплексных чисел и геометрический смысл вычисления обратного числа.

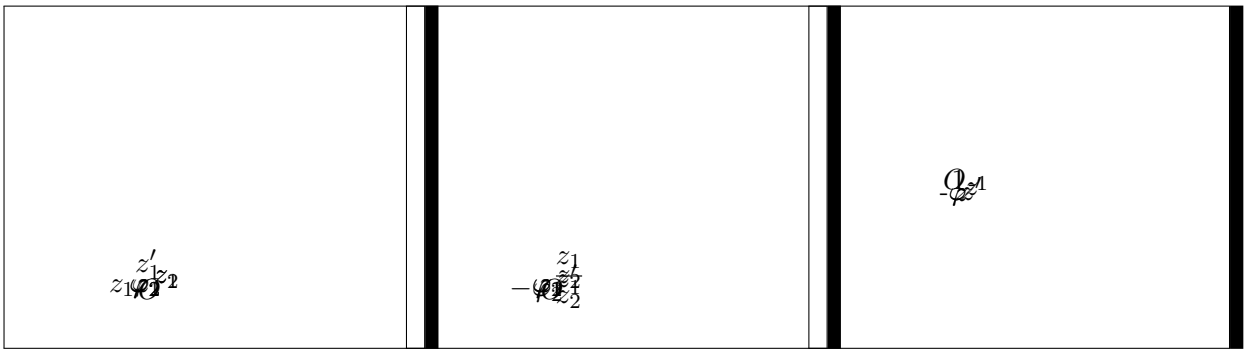

Рис. 12 Рис. 13 Рис. 14

Пусть заданы числа $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ и $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$, которые изображаются векторами на рис. 12–13.

1) Построение произведения $z_1 z_2$ проведено на рис.12. Вектор z_1 поворачиваем на угол φ_2 (с учётом знака φ_2) и получаем вектор $z_1' = r_1(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2))$. Затем вектор z_1' умножаем на $r_2 = |z_2|$ и получаем вектор, изображающий $r_1 r_2(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2)) = z_1 z_2$.

2) Построение частного $\dfrac{z_1}{z_2}$ проведено на рис.13. Вектор z_1 поворачиваем на угол $(-\varphi_2)$ и получаем вектор $z_1' = r_1(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$. Затем вектор z_1' делим на $r_2 = |z_2|$ и получаем вектор, изображающий $\dfrac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)) = \dfrac{z_1}{z_2}$.

3) На рис.14 для числа $z = r(\cos\varphi + i\sin\varphi) \neq 0$ проведено построение обратного числа $z^{-1} = \dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi))$. От вектора z, переходим к вектору $z' = \dfrac{1}{r}(\cos\varphi + i\sin\varphi)$. Затем от z' преобразованием симметрии относительно вещественной оси переходим к вектору $\dfrac{1}{r}(\cos(-\varphi) + i\sin(-\varphi)) = z^{-1}$. ■

Задача 2.3.1. *Запишите в тригонометрической форме следующие числа:* 1) i, 2) $-3i$, 3) -4, 4) 2, 5) $-4-4i$, 6) $-\sqrt{3}+i$, 7) $-3+2i$.

Решение. Для того, чтобы число $z = x+yi \neq 0$, где $x = \mathrm{Re}(z)$, $y = \mathrm{Im}(z)$ записать в тригонометрической форме $z = r(\cos\varphi + i\sin\varphi)$, где $r = |z|$, $\varphi = \arg(z)$, сначала вычисляем модуль $r = |z| = \sqrt{x^2+y^2}$, затем из равенств $\cos\varphi = \dfrac{x}{r}$, $\sin\varphi = \dfrac{y}{r}$ находим аргумент $\varphi = \arg(z)$.

1) Для $z = i = 0+1i$ имеем $x = \mathrm{Re}(z) = 0$, $y = \mathrm{Im}(z) = 1$, $r = |z| = \sqrt{x^2+y^2} = \sqrt{0^2+1^2} = 1$, $\cos\varphi = \dfrac{x}{r} = 0$, $\sin\varphi = \dfrac{y}{r} = 1$. Отсюда находим один из возможных аргументов $\varphi = \dfrac{\pi}{2}$ или все возможные аргументы $\varphi = \dfrac{\pi}{2} + 2\pi k$, где $k \in Z$.

Ответ: $i = \cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2}$ или $i = \cos\left(\dfrac{\pi}{2} + 2\pi k\right) + i\sin\left(\dfrac{\pi}{2} + 2\pi k\right)$, $k \in Z$.

2) Для $z = -3i = 0 + (-3)i$ имеем $x = 0$, $y = -3$, $r = |z| = \sqrt{0^2+(-3)^2} = 3$, $\cos\varphi = \dfrac{x}{r} = 0$, $\sin\varphi = \dfrac{y}{r} = -1$. Отсюда находим $\varphi = -\dfrac{\pi}{2}$.

Ответ: $-3i = 3\left(\cos\left(-\dfrac{\pi}{2}\right) + i\sin\left(-\dfrac{\pi}{2}\right)\right)$.

3) Для $z = -4 = -4 + 0i$ имеем $x = -4$, $y = 0$, $r = |z| = \sqrt{(-4)^2+0^2} = 4$, $\cos\varphi = \dfrac{x}{r} = -1$, $\sin\varphi = \dfrac{y}{r} = 0$. Отсюда находим $\varphi = \pi$.

Ответ: $-4 = 4(\cos\pi + i\sin\pi)$.

4) Для $z = 2 = 2 + 0i$ имеем $x = 2$, $y = 0$, $r = |z| = \sqrt{2^2+0^2} = 2$, $\cos\varphi = \dfrac{x}{r} = 1$, $\sin\varphi = \dfrac{y}{r} = 0$. Отсюда находим $\varphi = 0$.

Ответ: $2 = 2(\cos 0 + i\sin 0)$.

5) Для $z = -4-4i$ имеем $x = -4$, $y = -4$, $r = |z| = \sqrt{(-4)^2+(-4)^2} = 4\sqrt{2}$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{\sqrt{2}}{2}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{\sqrt{2}}{2}$. Отсюда находим $\varphi = -\dfrac{3\pi}{4}$.

Ответ: $-4-4i = 4\sqrt{2}\left(\cos\left(-\dfrac{3\pi}{4}\right) + i\sin\left(-\dfrac{3\pi}{4}\right)\right)$.

6) Для $z = -\sqrt{3}+i$ имеем $x = -\sqrt{3}$, $y = 1$, $r = |z| = \sqrt{(-\sqrt{3})^2+1^2} = 2$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{\sqrt{3}}{2}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{1}{2}$. Отсюда находим $\varphi = \dfrac{5\pi}{6}$.

Ответ: $-\sqrt{3}+i=2\left(\cos\left(\frac{5\pi}{6}\right)+i\sin\left(\frac{5\pi}{6}\right)\right)$.

7) Для $z=-3+2i$ имеем $x=-3$, $y=2$, $r=|z|=\sqrt{(-3)^2+2)^2}=$ $=13$, $\cos\varphi=\frac{x}{r}=-\frac{3}{\sqrt{13}}$, $\sin\varphi=\frac{y}{r}=\frac{2}{\sqrt{13}}$. Отсюда находим $\operatorname{tg}\varphi=\frac{\sin\varphi}{\cos\varphi}=-\frac{2}{3}$ и во второй четверти $\varphi=\operatorname{arctg}\left(-\frac{2}{3}\right)+\pi=\pi-\operatorname{arctg}\left(\frac{2}{3}\right)$.

Ответ: $-3+2i=\sqrt{13}\left(\cos\left(\pi-\operatorname{arctg}\left(\frac{2}{3}\right)\right)+i\sin\left(\pi-\operatorname{arctg}\left(\frac{2}{3}\right)\right)\right)$. ▌

2.4. Свойства модуля комплексных чисел. Свойства операции сопряжения

Определение 1.

Пусть задано комплексное число $z=x+yi$. Модулем z называется вещественное число $\sqrt{x^2+y^2}$. Оно обозначается $|z|$.

Комплексное число $\overline{z}=x-yi$ называется сопряжённым к числу z.

Преобразование $z\mapsto\overline{z}$ называется операцией сопряжения.

Замечание. Модуль комплексного числа $z=x+0i$, $x\in R$ равен $\sqrt{x^2}$ и он совпадает с модулем вещественного числа x.

Теорема 2.4.1 (Свойства модуля).

Справедливы следующие утверждения

$$1)\ \forall z\in C\quad |z|=|-z|=|\overline{z}|, \tag{9}$$

$$2)\ \forall z_1,\ z_2\in C\quad |z_1z_2|=|z_1||z_2|, \tag{10}$$

$$3)\ \forall z_1,\ z_2\in C,\ z_2\neq 0\quad \left|\frac{z_1}{z_2}\right|=\frac{|z_1|}{|z_2|}, \tag{11}$$

$$4)\ \forall z_1,\ z_2\in C\quad ||z_1|-|z_2||\leqslant|z_1\pm z_2|\leqslant|z_1|+|z_2|. \tag{12}$$

Доказательство.

1) Если $z=x+yi$, то значения

$|z|=|x+yi|=\sqrt{x^2+y^2}$,

$|-z|=|-x-yi|=\sqrt{(-x)^2+(-y)^2}=\sqrt{x^2+y^2}$,

$|\overline{z}|=|x-yi|=\sqrt{x^2+(-y)^2}=\sqrt{x^2+y^2}$

совпадают, т. е. выполнено (9).

2), 3) Из теоремы 2.3.1 следуют (10) и (11).

4) Первый способ доказательства следует из геометрического смысла сложения и вычитания комплексных чисел как векторов, см. рис. 8 на стр. 43.

Второй способ использует тригонометрическую форму чисел:

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1), \text{ где } r_1 = |z_1|,\ \varphi_1 = \arg(z_1),$$
$$z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2), \text{ где } r_2 = |z_2|,\ \varphi_2 = \arg(z_2).$$

Найдём $z_1 \pm z_2$:

$$z_1 \pm z_2 = r_1(\cos\varphi_1 + i\sin\varphi_1) \pm r_2(\cos\varphi_2 + i\sin\varphi_2) =$$
$$= (r_1\cos\varphi_1 \pm r_2\cos\varphi_2) + (r_1\sin\varphi_1 \pm r_2\sin\varphi_2)i.$$

Далее выразим $|z_1 \pm z_2|^2$:

$$|z_1 \pm z_2|^2 = (r_1\cos\varphi_1 \pm r_2\cos\varphi_2)^2 + (r_1\sin\varphi_1 \pm r_2\sin\varphi_2)^2 =$$
$$= r_1^2\cos^2\varphi_1 \pm 2r_1r_2\cos\varphi_1\cos\varphi_2 + r_2^2\cos^2\varphi_2 +$$
$$+ r_1^2\sin^2\varphi_1 \pm 2r_1r_2\sin\varphi_1\sin\varphi_2 + r_2^2\sin\varphi_2^2 =$$
$$= r_1^2(\cos^2\varphi_1 + \sin^2\varphi_1) \pm 2r_1r_2(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) +$$
$$+ r_2^2(\cos^2\varphi_2 + \sin\varphi_2^2) = r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2.$$

Затем в силу неравенства $-1 \leqslant \pm\cos(\varphi_1 - \varphi_2) \leqslant 1$ получим

$$r_1^2 - 2r_1r_2 + r_2^2 \leqslant r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2 \leqslant r_1^2 + 2r_1r_2 + r_2^2, \text{ т. е.}$$
$$(r_1 - r_2)^2 \leqslant r_1^2 \pm 2r_1r_2\cos(\varphi_1 - \varphi_2) + r_2^2 \leqslant (r_1 + r_2)^2 \text{ или}$$
$$|r_1 - r_2|^2 \leqslant |z_1 \pm z_2|^2 \leqslant |r_1 + r_2|^2.$$

Отсюда следует неравенство $|r_1 - r_2| \leqslant |z_1 \pm z_2| \leqslant |r_1 + r_2|$, совпадающее с (12).

Теорема 2.4.2 (Свойства операции сопряжения).

Справедливы следующие утверждения.

$$1)\ \forall z = x + yi \in C \quad \begin{cases} z + \overline{z} = 2x = 2\mathrm{R}e(z), \\ z\overline{z} = x^2 + y^2 = |z|^2; \end{cases} \tag{13}$$

$$2)\ \forall z \in C \quad \overline{\overline{z}} = z, \tag{14}$$

$$3)\ \forall z \in C \quad z = \overline{z} \Leftrightarrow z \in R \Leftrightarrow \mathrm{I}m(z) = 0, \tag{15}$$

$$4)\forall z_1,\ z_2 \in C \quad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \tag{16}$$

$$5)\forall z_1,\ z_2 \in C \quad \overline{z_1 - z_2} = \overline{z_1} - \overline{z_2}, \tag{17}$$

$$6)\forall z_1,\ z_2 \in C \quad \overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}, \tag{18}$$

$$7)\forall z_1,\ z_2 \neq 0 \in C \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}. \tag{19}$$

$$\tag{20}$$

Доказательство. Пусть $z = x + yi$, тогда $\overline{z} = x - yi$.

1) $z + \overline{z} = 2x = 2\mathrm{R}e(z)$;

$z\overline{z} = (x + yi)(x - yi) = x^2 + y^2 = \left(\sqrt{x^2 + y^2}\right)^2 = |z|^2$.

2) $\overline{\overline{z}} = \overline{x - yi} = x + yi = z$.

3) $z = \overline{z} \Leftrightarrow x + yi = x - yi \Leftrightarrow 2yi = 0 \Leftrightarrow y = 0 \Leftrightarrow \mathrm{I}m(z) = 0$.

Пусть $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$, тогда $\overline{z_1} = x_1 - y_1 i$, $\overline{z_2} = x_2 - y_2 i$,

4) $z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i$,

$\overline{z_1} + \overline{z_2} = x_1 - y_1 i + x_2 - y_2 i = (x_1 + x_2) - (y_1 + y_2)i = \overline{z_1 + z_2}$.

5) $z_1 - z_2 = (x_1 - x_2) + (y_1 - y_2)i$,

$\overline{z_1} - \overline{z_2} = x_1 - y_1 i - (x_2 - y_2 i) = (x_1 - x_2) - (y_1 - y_2)i = \overline{z_1 - z_2}$.

6) $z_1 z_2 = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i$,

$\overline{z_1} \cdot \overline{z_2} = (x_1 x_2 - y_1 y_2) - (x_1 y_2 + y_1 x_2)i = \overline{z_1 z_2}$.

7) Обозначим $a = \dfrac{1}{|z_2|^2}$, $a \in R$.

$\dfrac{z_1}{z_2} = \dfrac{z_1 \overline{z_2}}{z_2 \overline{z_2}} = \dfrac{z_1 \overline{z_2}}{|z_2|^2} = z_1 \overline{z_2} a$, $\overline{\left(\dfrac{z_1}{z_2}\right)} = \overline{z_1 \overline{z_2} a} = \overline{z_1}\,\overline{\overline{z_2}}\,\overline{a} = \overline{z_1} z_2 a$,

$\dfrac{\overline{z_1}}{\overline{z_2}} = \dfrac{\overline{z_1}\, z_2}{\overline{z_2} z_2} = \dfrac{\overline{z_1} z_2}{|z_2|^2} = \overline{z_1} z_2 a$. ■

2.5. Возведение в степень. Извлечение корня

Определение 1. *Пусть z — комплексное число и n — целое число. Определим степени z^n следующим образом.*

$$z^n = \begin{cases} \underbrace{zz\ldots z}_{n} & \textit{при } n \geqslant 1, \\ 1 \quad \textit{при } n = 0,\ z \neq 0, \\ \left(\dfrac{1}{z}\right)^{-n} & \textit{при } n \leqslant -1,\ z \neq 0. \ \blacksquare \end{cases}$$

Из определения следуют привычные свойства степеней

$$\begin{aligned} &\forall n \in Z,\ z_1, z_2 \in C,\ z_1 \neq 0,\ z_2 \neq 0 \quad (z_1 z_2)^n = z_1^n z_2^n, \\ &\forall k \in N,\ n \in Z,\ z_1, z_2, \ldots, z_k \in C,\ z_1 \neq 0,\ z_2 \neq 0, \ldots, z_k \neq 0 \\ &\qquad (z_1 z_2 \ldots z_k)^n = z_1^n z_2^n \ldots z_k^n, \\ &\forall m, n \in Z,\ z \in C,\ z \neq 0 \quad z^m z^n = z^{m+n}, \\ &\forall m, n \in Z,\ z \in C,\ z \neq 0 \quad (z^m)^n = z^{mn}. \ \blacksquare \end{aligned} \tag{21}$$

I. Степень числа, заданного в алгебраической форме.

Найдём i^k для $k \geqslant 0$: $i^0 = 1$, $i^1 = i$, $i^2 = i^1 i = -1$, $i^3 = i^2 i = -1i = -i$, $i^4 = i^3 i = -ii = -i^2 = 1$, $i^5 = i^4 i = 1i = i$, $\ldots$. Понятно, что значения 1, i, -1, $-i$ повторяются:

$$i^{4m} = 1,\ i^{4m+1} = i,\ i^{4m+2} = -1,\ i^{4m+3} = -i. \tag{22}$$

Для возведения числа $z = x+yi$ в натуральную степень для небольшого показателя можно применить к выражению $(x+yi)^n$ формулу бинома Ньютона (она справедлива, т. к. её доказательство основано на коммутативности и ассоциативности сложения и умножения, а также на дистрибутивности умножения), а затем воспользоваться формулами (22). Предварительно можно представить степень с использованием формул (21). Итак, $(x+yi)^n = \sum\limits_{k=0}^{n} C_n^k x^{n-k}(yi)^k = \sum\limits_{k=0}^{n} \dfrac{n!}{k!(n-k)!} x^{n-k} y^k i^k = \sum\limits_{k=0}^{n} \dfrac{n(n-1)\ldots(n-(k+}{k!}$ $x^n + nx^{n-1}yi - \dfrac{n(n-1)}{2}x^{n-2}y^2 - \dfrac{n(n-1)(n-2)}{6}x^{n-3}y^3 i + \ldots + nxy^{n-1}i^{n-1} + y^n i^n$.

Для возведения числа $u+vi$ в отрицательную степень n найдём обратное число $\dfrac{1}{u+vi} = \dfrac{u-vi}{(u+vi)(u-vi)} = \dfrac{u-vi}{u^2+v^2} = \dfrac{u}{u^2+v^2} + \dfrac{-v}{u^2+v^2} i = x + yi$ и возведём полученное число в положительную степень $(x+yi)^{-n}$ так, как было указано выше. $\blacksquare$

II. Степень числа, заданного в тригонометрической форме.

Теорема 2.5.1 (Формула Муавра).

Для произвольного ненулевого комплексного числа $z = r(\cos\varphi + i\sin\varphi)$, $(r = |z|,\ \varphi = \arg(z))$ *при всех* $n \in Z$ *справедлива формула*

$$(r(\cos\varphi + i\sin\varphi))^n = r^n(\cos(n\varphi) + i\sin(n\varphi)), \tag{23}$$

то есть при возведении числа в степень n *его модуль возводится в степень* n, *а его аргумент умножается на* n.

Доказательство. При $n = 0$ имеем $(r(\cos\varphi + i\sin\varphi))^0 = 1$ и $r^0(\cos(0\varphi) + i\sin(0\varphi)) = 1$, т.е. формула верна. Для $n \in N$ получим утверждение из теоремы 2.3.1 об умножении комплексных чисел в тригонометрической форме. Возьмём в формуле (8) равные числа $z_1 = z_2 = \ldots = z_n = z = r(\cos\varphi + i\sin\varphi)$ и получим $z^n = r^n(\cos(n\varphi) + i\sin(n\varphi))$, т.е. формула (23) верна.

Далее при $n < 0$ имеем $(r(\cos\varphi + i\sin\varphi))^n = \dfrac{1}{(r(\cos\varphi + i\sin\varphi))^{-n}} = \dfrac{1}{r^{-n}(\cos(-n\varphi) + i\sin(-n\varphi))} = r^n(\cos(n\varphi) + i\sin(n\varphi))$, т. е. формула (23) верна. ■

III. Корень n-ой степени из комплексного числа.

Определение 2. *Пусть* $n \in N$. **Корнем n-ой степени** *из комплексного числа* z *называется комплексное число* w *такое, что* $w^n = z$.

Существует единственный корень n-ой степени из $z = 0$ — это $w = 0$, так как $w^n = 0 \Leftrightarrow |w^n| = 0 \Leftrightarrow |w|^n = 0 \Leftrightarrow |w| = 0 \Leftrightarrow w = 0$.

Получим формулу для нахождения корней n-ой степени из комплексного числа, заданного в тригонометрической форме.

Теорема 2.5.2 (Корень n-ой степени из комплексного числа).

Для любого числа $z = r(\cos\varphi + i\sin\varphi) \neq 0$, $(r = |z|,\ \varphi = \arg(z))$ *существует ровно* n *различных корней* n*-ой степени из* z:

$$w_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right),\ k = \overline{0, n-1}. \tag{24}$$

Доказательство. Пусть $w = \rho(\cos\alpha + i\sin\alpha)$ является корнем n-ой степени из $z = r(\cos\varphi + i\sin\varphi)$. Тогда $w^n = z \Rightarrow (\rho(\cos\alpha + i\sin\alpha))^n = r(\cos\varphi + i\sin\varphi) \Rightarrow \rho^n(\cos(n\alpha) + i\sin(n\alpha)) = r(\cos\varphi + i\sin\varphi) \Rightarrow \rho^n = r$, $n\alpha = \varphi + 2\pi k$, $k \in Z \Rightarrow \rho = \sqrt[n]{r}$, $\alpha = \dfrac{\varphi + 2\pi k}{n}$, $k \in Z$. Отсюда получаем, что w имеет вид $w = \sqrt[n]{r}\left(\cos\dfrac{\varphi + 2\pi k}{n} + i\sin\dfrac{\varphi + 2\pi k}{n}\right)$, $k \in Z$.

Обратно, если мы берём число $w = \sqrt[n]{r}\left(\cos\dfrac{\varphi + 2\pi k}{n} + i\sin\dfrac{\varphi + 2\pi k}{n}\right)$ при любом $k \in Z$, то $w^n = (\sqrt[n]{r})^n\left(\cos\dfrac{(\varphi + 2\pi k)n}{n} + i\sin\dfrac{(\varphi + 2\pi k)n}{n}\right) \Rightarrow w^n = r(\cos\varphi + i\sin\varphi) \Rightarrow w^n = z$, т. е. w является корнем n-ой степени из z.

Следовательно, только числа $w_k = \sqrt[n]{r}\left(\cos\dfrac{\varphi + 2\pi k}{n} + i\sin\dfrac{\varphi + 2\pi k}{n}\right)$, $k \in Z$ являются корнями n-ой степени из z.

Все w_k имеют модуль $\sqrt[n]{r}$ и на комплексной плоскости они расположены на окружности радиуса $\sqrt[n]{r}$ с центром в 0, аргумент w_k равен $\alpha_k = \dfrac{\varphi + 2\pi k}{n} = \dfrac{\varphi}{n} + \dfrac{2\pi}{n}k$ или отличается от α_k на $2\pi l$, $l \in Z$. Увеличение k на 1 влечёт увеличение α_k на $\dfrac{2\pi}{n}$. Взяв для k возрастающую последовательность значений от 0 до $n-1$, мы получим возрастающую последовательность n различных значений аргументов $\alpha_0 = \dfrac{\varphi}{n} < \alpha_1 = \dfrac{\varphi}{n} + \dfrac{2\pi}{n} < \alpha_2 = \dfrac{\varphi}{n} + \dfrac{2\pi \cdot 2}{n} < \ldots < \alpha_{n-1} = \dfrac{\varphi}{n} + \dfrac{2\pi \cdot (n-1)}{n}$, которые различаются между собой самое большое на $\dfrac{2\pi \cdot (n-1)}{n} < 2\pi$, т. е. принадлежат промежутку $\left[\dfrac{\varphi}{n}, \dfrac{\varphi}{n} + 2\pi\right)$. Это означает, что на комплексной плоскости числа w_k, $k = \overline{0, n-1}$ расположены на окружности радиуса $\sqrt[n]{r}$ с центром в 0 и они делят эту окружность на n равных частей, поскольку разность аргументов $\alpha_{k+1} - \alpha_k = \dfrac{2\pi}{n}$.

Каждое число w_k для $k \in Z$ совпадает с одним из корней $\{w_0, w_1, w_2, \ldots, w_{n-1}\}$. Действительно, при $k = nm + r$ $(m, r \in Z,\ 0 \leqslant r \leqslant n-1)$ имеем

$$\frac{\varphi + 2\pi k}{n} = \frac{\varphi + 2\pi nm + 2\pi r}{n} = \frac{\varphi + 2\pi r}{n} + 2\pi m.$$

Следовательно, $w_k = w_r \in \{w_0,\, w_1,\, w_2, \ldots, w_{n-1}\}$.

Следствия. 1. Формулу (24) можно записать в виде равенства двух множеств

$$\sqrt[n]{r(\cos\varphi + i\sin\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi k}{n} + i\sin\frac{\varphi+2\pi k}{n}\right),\ k = \overline{0, n-1}. \tag{25}$$

2. Все значения $\{w_0,\, w_1,\, w_2, \ldots, w_{n-1}\}$ корня n-ой степени из числа $z \neq 0$ на комплексной плоскости расположены на окружности радиуса $\sqrt[n]{|z|}$ с центром в нуле и они являются вершинами правильного n-угольника, причём $\arg(w_0) = \dfrac{\arg(z)}{n}$ и $\arg(w_k) = \arg(w_{k-1}) + \dfrac{2\pi}{n}$, $k = \overline{1, n-1}$.

3. Все значения корня n-ой степени из $z \neq 0$ можно получить умножением одного из его значений на все значения корня n-ой степени из единицы $\varepsilon_m = \cos\dfrac{2\pi m}{n} + i\sin\dfrac{2\pi m}{n}$, $m = \overline{0, n-1}$.

Доказательство. 3. Числа $w_k = \sqrt[n]{r}\left(\cos\dfrac{\varphi+2\pi k}{n} + i\sin\dfrac{\varphi+2\pi k}{n}\right)$, $k \in Z$ дают все значения корня n-ой степени из числа $z = r(\cos\varphi + i\sin\varphi)$.

Возьмём одно значение w_p и умножим его на все значения корня n-ой степени из единицы $\varepsilon_m = \cos\dfrac{2\pi m}{n} + i\sin\dfrac{2\pi m}{n}$, $m \in Z$. Получим

$$w_p\varepsilon_m = \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi p}{n} + i\sin\frac{\varphi+2\pi p}{n}\right)\left(\cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}\right) =$$

$$= \sqrt[n]{r}\left(\cos\frac{\varphi+2\pi(p+m)}{n} + i\sin\frac{\varphi+2\pi(p+m)}{n}\right).$$ Поскольку $p+m$ принимает любые целые значения последнее выражение содержит все значения корней w_k, $k = \overline{0, n-1}$. ■

IV. Квадратный корень из комплексного числа в алгебраической форме.

Из предыдущего раздела следует существование двух значений квадратного корня из заданного ненулевого комплексного числа $z = a + bi$, записанного в тригонометрической форме. Получим эти значения чисто в алгебраической форме $x + yi$, где x и y определяются с помощью радикалов.

Пусть $\sqrt{a+bi} = x+yi$, $x, y \in R$. Это означает, что $(x+yi)^2 = a+bi \Leftrightarrow x^2+2xyi-y^2 = a+bi \Leftrightarrow (x^2-y^2)+2xyi = a+bi \Leftrightarrow$

$$\begin{cases} x^2-y^2 = a, \\ 2xy = b. \end{cases} \tag{26}$$

Возводя в квадрат обе части каждого из уравнений системы (26), а затем складывая их, получим уравнение, являющееся следствием системы (26):

$(x^2-y^2)^2+4x^2y^2 = a^2+b^2 \Leftrightarrow x^4-2x^2y^2+y^4+4x^2y^2 = a^2+b^2 \Leftrightarrow$
$\Leftrightarrow x^4+2x^2y^2+y^4 = a^2+b^2 \Leftrightarrow (x^2+y^2)^2 = a^2+b^2 \Leftrightarrow x^2+y^2 = +\sqrt{a^2+b^2}$.
Корень квадратный взят со знаком плюс, т.к. $x^2+y^2 \in R$ и $x^2+y^2 \geqslant 0$.

Добавив уравнение $x^2+y^2 = \sqrt{a^2+b^2}$ к (26), получим эквивалентную систему

$$\begin{cases} x^2+y^2 = \sqrt{a^2+b^2}, \\ x^2-y^2 = a, \\ 2xy = b. \end{cases} \tag{27}$$

Складывая и вычитая два уравнения в (27), придём к эквивалентной системе

$$\begin{cases} x^2 = \dfrac{a+\sqrt{a^2+b^2}}{2}, \\ y^2 = \dfrac{-a+\sqrt{a^2+b^2}}{2}, \\ 2xy = b. \end{cases} \tag{28}$$

При $b=0$, $a>0$ из системы (28) найдём $x^2=a$, $y^2=0$ и получим $x=\pm\sqrt{a}$, $y=0$. Следовательно,

$$\sqrt{a+0i} = \pm\sqrt{a}+0i \text{ при } a>0.$$

При $b=0$, $a<0$ из системы (28) найдём $x^2=0$, $y^2=-a>0$ и получим $x=0$, $y=\pm\sqrt{-a}$. Следовательно,

$$\sqrt{a+0i} = 0\pm\sqrt{-a}i \text{ при } a<0.$$

Далее рассмотрим случай $b \neq 0$.

Для любых $a, b \in R$, $b \neq 0$ справедливы неравенства

$$a^2+b^2 > a^2 \Leftrightarrow \sqrt{a^2+b^2} > |a| \Leftrightarrow \begin{cases} \sqrt{a^2+b^2} > a, \\ \sqrt{a^2+b^2} > -a, \end{cases} \Leftrightarrow \begin{cases} -a+\sqrt{a^2+b^2} > 0, \\ a+\sqrt{a^2+b^2} > 0. \end{cases}$$

Тогда из системы (28) найдём

$$x = \pm\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}, \quad y = \pm\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}. \tag{29}$$

Причём в силу уравнения $2xy = b$ знаки для x и y в (29) выбираем так, чтобы знак произведения xy совпал со знаком b. При $b > 0$ получаем две пары значений $x > 0$, $y > 0$ и $x < 0$, $y < 0$. При $b < 0$ получаем две пары значений $x > 0$, $y < 0$ и $x < 0$, $y > 0$. Следовательно,

$$\sqrt{a+bi} = \pm\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} + \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\, i\right) \quad \text{при } b > 0, \tag{30}$$

$$\sqrt{a+bi} = \pm\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} - \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\, i\right) \quad \text{при } b < 0. \tag{31}$$

Обе формулы (30), (31) дают $\pm\sqrt{a}$ при $b = 0$, $a > 0$, $\pm\sqrt{-a}i$ при $b = 0$, $a < 0$.

Для проверки нужно возвести правые части формул (30) или (31) в квадрат и получить заданное значение $a + bi$.

Итак, здесь ещё раз показано, что существует два значения квадратного корня из каждого ненулевого комплексного числа. Эти значения выражаются через радикалы по формулам (30) или (31) и являются противоположными числами. ■

Задача 2.5.1. *Найдите* $(1-i)^{20}$.

Решение. Преобразуем $w = (1-i)^{20} = \left((1-i)^2\right)^{10} = (1-2i+i^2)^{10} = (-2i)^{10} = (-2)^{10}i^{10}$. Поскольку $(-2)^{10} = 2^{10} = 1024$ и $i^{10} = i^2 = -1$, получим $w = -1024$.

Задача 2.5.2. *Найдите* $(-3+i)^{-4}$.

Решение. Первый способ. Представим отрицательную степень в виде положительной степени обратного числа $(-3+i)^{-4} = \left(\dfrac{1}{-3+i}\right)^4$. Затем выполним деление $\dfrac{1}{-3+i} = \dfrac{-3-i}{(-3+i)(-3-i)} = \dfrac{-3-i}{(-3)^2-i^2} = \dfrac{-3-i}{10}$. Далее возведём в степень $\left(\dfrac{-3-i}{10}\right)^4 = \dfrac{(-1)^4\left((3+i)^2\right)^2}{10^4} = \dfrac{(9+6i-1)^2}{10^4} = \dfrac{(8+6i)^2}{10^4} = \dfrac{(2(4+3i))^2}{10^4} = \dfrac{2^2(4+3i)^2}{10^4} = \dfrac{4(16+24i-9)}{10^4} = \dfrac{7+24i}{2500} = \dfrac{7}{2500} + \dfrac{6}{625}i$.

Второй способ. Представим отрицательную степень в виде обратного значения к положительной степени $(-3+i)^{-4} = \dfrac{1}{(-3+i)^4}$. Затем возведём в положительную степень $(-3+i)^4 = ((-3+i)^2)^2 = (9-6i-1)^2 = (8-6i)^2 = (2(4-3i))^2 = 4(16-24i-9) = 4(7-24i)$. Далее выполним деление $\dfrac{1}{(-3+i)^4} = \dfrac{1}{4(7-24i)} = \dfrac{7+24i}{4(7-24i)(7+24i)} = \dfrac{7+24i}{4(49+576)} = \dfrac{7+24i}{4\cdot 625} = \dfrac{7}{2500} + \dfrac{6}{625}i$.

Задача 2.5.3. *Найдите* $(-1+\sqrt{3}i)^{22}$.

Решение. Представим число $z = -1+\sqrt{3}i$ в тригонометрической форме. Находим $x = \mathrm{Re}(z) = -1$, $y = \mathrm{Im}(z) = \sqrt{3}$, $r = |z| = \sqrt{x^2+y^2} = \sqrt{1+3} = \sqrt{4} = 2$, $\cos\varphi = \dfrac{x}{r} = -\dfrac{1}{2}$, $\sin\varphi = \dfrac{y}{r} = \dfrac{\sqrt{3}}{2}$, $\varphi = \dfrac{2\pi}{3}$. Отсюда $-1+\sqrt{3}i = 2\left(\cos\dfrac{2\pi}{3} + i\sin\dfrac{2\pi}{3}\right)$. Далее $(-1+\sqrt{3}i)^{22} = \left(2\left(\cos\dfrac{2\pi}{3} + i\sin\dfrac{2\pi}{3}\right)\right)^{22} = 2^{22}\left(\cos\dfrac{2\pi\cdot 22}{3} + i\sin\dfrac{2\pi\cdot 22}{3}\right) = 2^{22}\left(\cos\dfrac{42\pi+2\pi}{3} + i\,\mathrm{s}\right.$ $= 2^{22}\left(\cos\left(14\pi+\dfrac{2\pi}{3}\right) + i\sin\left(14\pi+\dfrac{2\pi}{3}\right)\right) = 2^{22}\left(\cos\left(\dfrac{2\pi}{3}\right) + i\sin\left(\dfrac{2\pi}{3}\right)\right) = 2^{22}\left(-\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}\right) = 2^{21}(-1+\sqrt{3}i)$.

Задача 2.5.4. *Найдите* $\left(\dfrac{4}{5} - \dfrac{3}{5}i\right)^{-100}$.

Решение. Представим число $z = \dfrac{4}{5} - \dfrac{3}{5}i$ в тригонометрической

форме. Находим $x = \mathrm{Re}(z) = \dfrac{4}{5}$, $y = \mathrm{Im}(z) = -\dfrac{3}{5}$, $r = |z| = \sqrt{x^2+y^2} =$
$= \sqrt{\left(\dfrac{4}{5}\right)^2 + \left(\dfrac{3}{5}\right)^2} = \sqrt{\dfrac{16}{25} + \dfrac{9}{25}} = 1$, $\cos\varphi = \dfrac{x}{r} = \dfrac{4}{5}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{3}{5}$.
Тогда $\operatorname{tg}\varphi = -\dfrac{3}{4}$ и в четвёртой четверти возьмём $\varphi = \operatorname{arctg}\left(-\dfrac{3}{4}\right) =$
$-\operatorname{arctg}\dfrac{3}{4}$.
Отсюда $\dfrac{4}{5} - \dfrac{3}{5}i = \cos\left(-\operatorname{arctg}\dfrac{3}{4}\right) + i\sin\left(-\operatorname{arctg}\dfrac{3}{4}\right)$. И далее по формуле Муавра находим $\left(\dfrac{4}{5} - \dfrac{3}{5}i\right)^{-100} = \left(\cos\left(-\operatorname{arctg}\dfrac{3}{4}\right) + i\sin\left(-\operatorname{arctg}\dfrac{3}{4}\right)\right)^{-100} =$
$= \cos\left(100\operatorname{arctg}\dfrac{3}{4}\right) + i\sin\left(100\operatorname{arctg}\dfrac{3}{4}\right) \approx 0,053 + 0,999i$.

Задача 2.5.5. *Выполнить действия* $\dfrac{(-\sqrt{3}-i)^{20}}{(1-\sqrt{3}i)^4}$.

Решение.

Представим числа $-\sqrt{3}-i$, $1-\sqrt{3}i$ в тригонометрической форме и выполним действия.

Для $z = -\sqrt{3}-i$ имеем $x = \mathrm{Re}(z) = -\sqrt{3}$, $y = \mathrm{Im}(z) = -1$, тогда $r = |z| = \sqrt{x^2+y^2} = \sqrt{(-\sqrt{3})^2 + (-1)^2} = \sqrt{4} = 2$, $\cos\varphi = \dfrac{x}{r} =$
$-\dfrac{\sqrt{3}}{2}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{1}{2}$. Отсюда $\varphi = \dfrac{7\pi}{6}$, $-\sqrt{3}-i = 2\left(\cos\dfrac{7\pi}{6} + i\sin\dfrac{7\pi}{6}\right)$ и $(-\sqrt{3}-$
$i)^{20} = 2^{20}\left(\cos\dfrac{7\pi\cdot 20}{6} + i\sin\dfrac{7\pi\cdot 20}{6}\right) = 2^{20}\left(\cos\dfrac{70\pi}{3} + i\sin\dfrac{70\pi}{3}\right)$.

Для $z = 1-\sqrt{3}i$ имеем $x = \mathrm{Re}(z) = 1$, $y = \mathrm{Im}(z) = -\sqrt{3}$, тогда $r = |z| = \sqrt{1+3} = 2$, $\cos\varphi = \dfrac{x}{r} = \dfrac{1}{2}$, $\sin\varphi = \dfrac{y}{r} = -\dfrac{\sqrt{3}}{2}$. Отсюда $\varphi = -\dfrac{\pi}{3}$, $1-\sqrt{3}i = 2\left(\cos\left(-\dfrac{\pi}{3}\right) + i\sin\left(-\dfrac{\pi}{3}\right)\right)$ и $(1-\sqrt{3}i)^4 =$
$= 2^4\left(\cos\left(-\dfrac{4\pi}{3}\right) + i\sin\left(-\dfrac{4\pi}{3}\right)\right)$. Далее выполняем деление $\dfrac{(-\sqrt{3}-i)^{20}}{(1-\sqrt{3}i)^4} =$

$$\frac{2^{20}\left(\cos\dfrac{70\pi}{3} + i\sin\dfrac{70\pi}{3}\right)}{2^4\left(\cos\left(-\dfrac{4\pi}{3}\right) + i\sin\left(-\dfrac{4\pi}{3}\right)\right)} = \frac{2^{20}}{2^4}\left(\cos\frac{74\pi}{3} + i\sin\frac{74\pi}{3}\right) = 2^{16}\left(\cos\frac{72\pi+2\pi}{3} + \right.$$

$$= 2^{16}\left(\cos\left(24\pi + \frac{2\pi}{3}\right) + i\sin\left(24\pi + \frac{2\pi}{3}\right)\right) = 2^{16}\left(\cos\left(\frac{2\pi}{3}\right) + i\sin\left(\frac{2\pi}{3}\right)\right) =$$
$$= 2^{16}\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) = 2^{15}(-1+\sqrt{3}i).$$

Задача 2.5.6. *Найдите* $\sqrt{3+4i}$.

Решение. Обозначим $\sqrt{3+4i} = x+yi$. Тогда $(x+yi)^2 = 3+4i \Leftrightarrow$

$$(x^2-y^2)+2xyi = 3+4i \Leftrightarrow \begin{cases} x^2-y^2=3, \\ 2xy=4, \end{cases} \Leftrightarrow \begin{cases} (x^2-y^2)^2+4x^2y^2 = 3^2+4^2, \\ x^2-y^2=3, \\ 2xy=4, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} (x^2+y^2)^2 = 25, \\ x^2-y^2=3, \\ 2xy=4, \end{cases} \Leftrightarrow \begin{cases} x^2+y^2=+5, \\ x^2-y^2=3, \\ xy=2, \end{cases} \Leftrightarrow \begin{cases} 2x^2=8, \\ 2y^2=2, \\ xy=2, \end{cases} \Leftrightarrow \begin{cases} x^2=4, \\ y^2=1, \\ xy=2, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} x=\pm 2, \\ y=\pm 1, \\ xy>0, \end{cases} \Leftrightarrow x+yi = \pm(2+i).$$ Ответ: $2+i$, $-2-i$.

Задача 2.5.7. *Найдите* $\sqrt[4]{-1}$.

Решение. Поскольку $w = \sqrt[4]{z} \Leftrightarrow w^4 = z \Leftrightarrow (w^2)^2 = z \Leftrightarrow w^2 = \sqrt{z} \Leftrightarrow w = \sqrt{\sqrt{z}}$. Следовательно, нужно найти $\sqrt{\sqrt{-1}}$.

В силу $i^2 = -1$ имеем $\sqrt{-1} = \pm i$. Далее обозначим $\sqrt{\pm i} = x+yi$.

Тогда $(x+yi)^2 = \pm i \Leftrightarrow (x^2-y^2)+2xyi = \pm i \Leftrightarrow \begin{cases} x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow$

$$\Leftrightarrow \begin{cases} (x^2-y^2)^2+4x^2y^2 = 0+1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow \begin{cases} (x^2+y^2)^2 = 1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow \begin{cases} x^2+y^2=1, \\ x^2-y^2=0, \\ 2xy=\pm 1, \end{cases} \Leftrightarrow$$

$$\Leftrightarrow\Leftrightarrow \begin{cases} x^2=1/2, \\ y^2=1/2, \\ xy=\pm 1/2, \end{cases} \Leftrightarrow \begin{cases} x=\pm 1/\sqrt{2}, \\ y=\pm 1/\sqrt{2}, \\ xy=\pm 1/2. \end{cases}$$ Отсюда находим $x+yi = \pm\frac{\sqrt{2}}{2}(1\pm i)$.

Ответ: $\dfrac{\sqrt{2}}{2}(1+i)$, $-\dfrac{\sqrt{2}}{2}(1+i)$, $\dfrac{\sqrt{2}}{2}(1-i)$, $-\dfrac{\sqrt{2}}{2}(1-i)$.

Задача 2.5.8. *Найдите* $\sqrt[4]{-8-8\sqrt{3}i}$ *и изобразите результаты на комплексной плоскости.*

Решение. Представим число $z=-8-8\sqrt{3}i$ в тригонометрической форме и извлечём корень.

Для $z=-8-8\sqrt{3}i$ имеем $x=\mathrm{Re}(z)=-8$, $y=\mathrm{Im}(z)=-8\sqrt{3}$, тогда $r=|z|=\sqrt{x^2+y^2}=\sqrt{8^2+8^2\cdot 3}=\sqrt{8^2\cdot 4}=16$, $\cos\varphi=\dfrac{x}{r}=-\dfrac{1}{2}$, $\sin\varphi=\dfrac{y}{r}=-\dfrac{\sqrt{3}}{2}$, $\varphi=\dfrac{4\pi}{3}$. Отсюда $-8-8\sqrt{3}i=16\left(\cos\dfrac{4\pi}{3}+i\sin\dfrac{4\pi}{3}\right)$ и далее $\sqrt[4]{16\left(\cos\dfrac{4\pi}{3}+i\sin\dfrac{4\pi}{3}\right)}=\sqrt[4]{16}\left(\cos\dfrac{\frac{4\pi}{3}+2\pi k}{4}+i\sin\dfrac{\frac{4\pi}{3}+2\pi k}{4}\right)=$ $=2\left(\cos\left(\dfrac{\pi}{3}+\dfrac{\pi k}{2}\right)+i\sin\left(\dfrac{\pi}{3}+\dfrac{\pi k}{2}\right)\right)$, $k=\overline{0,3}$.

Обозначим полученные значения через w_k, тогда ответами в тригонометрической форме являются

$$w_0=2\left(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}\right),\qquad w_1=2\left(\cos\frac{5\pi}{6}+i\sin\frac{5\pi}{6}\right),$$
$$w_2=2\left(\cos\frac{4\pi}{3}+i\sin\frac{4\pi}{3}\right),\quad w_3=2\left(\cos\frac{11\pi}{6}+i\sin\frac{11\pi}{6}\right).$$

Ответы изображены на рис.15, где точки, изображающие числа w_0, w_1, w_2, w_3, являются вершинами квадрата, вписанного в окружность радиуса 2 с центром в начале координат. Для w_0 имеем $\arg(w_0)=\arg(z)/4=\pi/3$ и $\arg(w_1)-\arg(w_0)=\arg(w_2)-\arg(w_1)=\arg(w_3)-\arg(w_2)=\arg(w_0)-\arg(w_3)=\pi/2$.

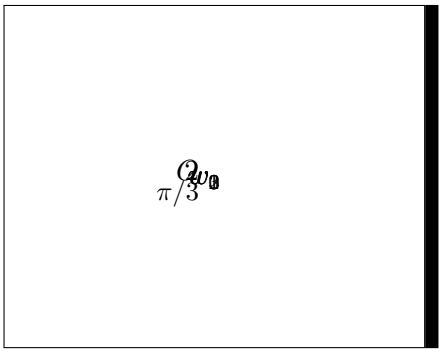

Рис. 15

3. Многочлены

3.1. Кольцо многочленов от одной переменной

Определение 1. *Пусть P — поле.*

Пусть задан счётный набор чисел из поля P:

$$f_0,\ f_1,\ f_2,\ \ldots,\ f_k,\ \ldots, \tag{32}$$

причём, начиная с некоторого номера $m \geqslant 0$, все $f_k = 0$.

Многочленом (полиномом) от переменной x над полем P с коэффициентами (32) называется выражение

$$f_0 + f_1\,x + f_2\,x^2 + \ldots + f_n\,x^n + \ldots \quad \text{или} \quad \sum_{k=0}^{\infty} f_k\,x^k. \tag{33}$$

Обозначение: f или $f(x)$.

Нулевым называем многочлен $\sum\limits_{k=0}^{\infty} 0\,x^k$ со всеми нулевыми коэффициентами. Обозначение: 0.

Через $P[x]$ обозначим множество всех многочленов от переменной x над полем P.

При записи многочлена $f = \sum\limits_{k=0}^{\infty} f_k\,x^k$ приняты некоторые соглашения:

- вместо $3+0x+\sqrt{2}x^2+0x^3+0x^4+\pi x^5+0x^6+\ldots = 3+0x+\sqrt{2}x^2+0x^3+0x^4+\pi x^5+\sum\limits_{k=6}^{\infty} 0\,x^k$ пишем $3+\sqrt{2}x^2+\pi x^5$ (опускаем все члены с нулевыми коэффициентами);
- вместо $(-2)+(-3)x+1x^2+0x^3+(-1)x^4+0x^5+\ldots$ пишем $-2-3x+x^2-x^4$ (свободный член $(-p)$, где $p>0$ меняем на $-p$; запись $+1x^k$, $k \geqslant 1$ меняем на $+x^k$; запись $+(-p)x^k$, $k \geqslant 1$, $p>0$ меняем на $-px^k$; запись $+(-1)x^k$, $k \geqslant 1$ меняем на $-x^k$).

Определение 2. *Многочлены $f(x) = \sum\limits_{k=0}^{\infty} f_k\,x^k$ и $g(x) = \sum\limits_{k=0}^{\infty} g_k\,x^k$ из $P[x]$ называются равными, если равны все их соответствующие коэффициенты $f_k = g_k$, $k = 0, 1, 2, \ldots$.*

Обозначение: равных многочленов $f(x) = g(x)$ или $f = g$, неравных многочленов $f(x) \neq g(x)$ или $f \neq g$.[1]

Определение 3. *Степенью ненулевого многочлена $f = \sum\limits_{k=0}^{\infty} f_k\,x^k$ называется максимальное число $n \in Z_+$, для которого коэффициент $f_n \neq 0$, (тогда f_n называют старшим коэффициентом).*

[1] Лучше было бы писать $f(x) \equiv g(x)$ и $f(x) \not\equiv g(x)$.

Степень нулевого многочлена считается равной символу $(-\infty)$.

Обозначение степени многочлена: $\deg(f)$ *или* $\deg f$.

Замечания.

1. Степень многочлен $f(x) = \sum\limits_{k=0}^{\infty} f_k\, x^k$ равна $n \in Z_+$ если $f_n \neq 0$ и все $f_k = 0$ при $k > n$.

2. Многочлен нулевой степени — это $f(x) = f_0$, причём $f_0 \neq 0$.

3. Многочлен степени n можно записать в виде
$f(x) = \sum\limits_{k=0}^{n} f_k\, x^k = f_0 + f_1\, x + f_2\, x^2 + \ldots + f_{n-1}\, x^{n-1} + f_n\, x^n$ (по возрастанию степеней),
$f(x) = \sum\limits_{k=0}^{n} a_k\, x^{n-k} = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$ (по убыванию степеней),
здесь $a_k = f_{n-k}$, $k = \overline{0, n}$.

4. Операции и неравенства с $(-\infty)$:
$(-\infty) + n = n + (-\infty) = (-\infty)$, если $n \in Z_+$;
$(-\infty) + (-\infty) = (-\infty)$;
$(-\infty) < n$, если $n \in Z_+$.

Введём на множестве многочленов $P[x]$ алгебраические операции.

Определение 4. *Пусть* $f \in P[x]$, $g \in P[x]$

$$f = f_0 + f_1\, x + f_2\, x^2 + \ldots + f_n\, x^n + \ldots = \sum_{k=0}^{\infty} f_k\, x^k,$$

$$g = g_0 + g_1\, x + g_2\, x^2 + \ldots + g_n\, x^n + \ldots = \sum_{k=0}^{\infty} g_k\, x^k.$$

I. Суммой многочленов f *и* g *называется выражение*

$$(f_0 + g_0) + (f_1 + g_1)\, x + \ldots + (f_k + g_k)\, x^k + \ldots = \sum_{k=0}^{\infty} (f_k + g_k)\, x^k. \quad (34)$$

Обозначение $f + g$.

II. Произведением многочленов f *и* g *называется выражение*[2]

$$f_0\, g_0 + (f_0\, g_1 + f_1\, g_0)\, x + (f_0\, g_2 + f_1\, g_1 + f_2\, g_0)\, x^2 + \ldots + (f_0\, g_k + f_1\, g_{k-1} + \ldots + f_{k-1}\, g_1 + f_k\, g_0)\, x^k + \ldots = \sum_{k=0}^{\infty} \Big(\sum_{r+s=k} f_r\, g_s \Big) x^k. \quad (35)$$

Обозначение fg.

[2] Суммирование в $\sum\limits_{r+s=k} f_r\, g_s$ ведётся по всем целым индексам $r \geqslant 0, s \geqslant 0$, для которых $r+s = k$.

Лемма 3.1.1. *Если f и g многочлены из $P[x]$, то сумма $f+g$ и произведение fg также являются многочленами из $P[x]$ и для них выполнены свойства:*

$$deg(f+g) \leqslant \max\{deg(f), deg(g)\}, \tag{36}$$

$$deg(fg) = deg(f) + deg(g). \tag{37}$$

Доказательство. Пусть $f \in P[x]$, $g \in P[x]$,

$$f = f_0 + f_1\,x + f_2\,x^2 + \ldots + f_k\,x^k + \ldots = \sum_{k=0}^{\infty} f_k\,x^k, \;\; m = deg(f),$$

$$g = g_0 + g_1\,x + g_2\,x^2 + \ldots + g_k\,x^k + \ldots = \sum_{k=0}^{\infty} g_k\,x^k, \;\; n = deg(g).$$

I. Рассмотрим

$$f+g = (f_0+g_0) + (f_1+g_1)\,x + (f_2+g_2)\,x^2 + \ldots + (f_k+g_k)\,x^k + \ldots.$$

При всех $k \in Z_+$ коэффициенты $f_k, g_k \in P$, поэтому $f_k + g_k \in P$. При $k > m$ все $f_k = 0$, при $k > n$ все $g_k = 0$, тогда все $f_k + g_k = 0$ при $k > \max\{m, n\}$. Следовательно, $f+g$ — многочлен из $P[x]$, а его степень не превосходит $\max\{m, n\}$, т. е. выполнено неравенство (36).

II. Рассмотрим

$$fg = f_0\,g_0 + (f_0\,g_1 + f_1\,g_0)\,x + (f_0\,g_2 + f_1\,g_1 + f_2\,g_0)\,x^2 + \ldots +$$
$$+(f_0\,g_k + f_1\,g_{k-1} + \ldots + f_{k-1}\,g_1 + f_k\,g_0)\,x^k + \ldots.$$

Коэффициенты f_r, $g_s \in P$ при всех $r, s \in Z_+$, поэтому при всех $k \in Z_+$ суммы $\sum\limits_{r+s=k} f_r\,g_s \in P$.

а) Если хотя бы один многочленов f и g равен 0, то $fg = 0 \in P[x]$, $\deg(fg) = -\infty$, $\deg(f) + \deg(g) = -\infty$ и равенство (37) выполнено.

б) Рассмотрим случай, когда $f \neq 0$ и $g \neq 0$.

При этом выполняются условия $m = \deg(f) \geqslant 0$ и $n = \deg(g) \geqslant 0$, коэффициенты $f_k = 0$ при $k > m$, $f_m \neq 0$ и $g_k = 0$ при $k > n$, $g_n \neq 0$.

Обозначим $h = fg$, где $h = \sum\limits_{k=0}^{\infty} h_k\,x^k$ и $h_k = \sum\limits_{r+s=k} f_r\,g_s$.

Докажем, что при $k > m+n$ коэффициенты $h_k = 0$, а при $k = m+$ $+\,n$ коэффициент $h_k = f_m\,g_n \neq 0$.

Рассмотрим $h_k = \sum\limits_{r+s=k} f_r\,g_s$ при $k > m+n$. Здесь $r+s = k$ и $k > m+n$, т. е. $r+s > m+n$. Из неравенства $r+s > m+n$ следует, что $r > m$ или $s > n$. Т. к. в противном случае $r \leqslant m$, $s \leqslant n$ и $r+s \leqslant m+n$.

Если $r > m$, то $f_r = 0$ и $f_r\, g_s = 0$. Если $s > n$, то $g_s = 0$ и $f_r\, g_s = 0$. Итак, при $k > m + n$ все произведения $f_r\, g_s = 0$ для $r + s = k$, и тогда $h_k = \sum\limits_{r+s=k} f_r\, g_s = 0$. Отсюда следует, что fg— многочлен из $P[x]$ и $deg(fg) \leqslant m + n$.

Рассмотрим $h_k = \sum\limits_{r+s=k} f_r\, g_s$ при $k = m+n$. Здесь $r+s = k = m+n$. Если $r > m$, то $f_r = 0$ и $f_r\, g_s = 0$. Если $r = m$, то $s = m - r + n = n$ и $f_r\, g_s = f_m\, g_n \neq 0$. Если $r < m \Leftrightarrow m - r > 0$, то $s = m - r + n > 0 + n = n \Rightarrow s > n$, следовательно, $g_s = 0$ и $f_r\, g_s = 0$. Итак, $h_{m+n} = f_m\, g_n \neq 0$. Отсюда следует, что $deg(fg) = m + n = deg(f) + deg(g)$, т.е. выполнено равенство (37).

Лемма доказана.

Следствия.

1. *Формула* $f + g$ *для* $f, g \in P[x]$ *задаёт отображение из* $P[x] \times P[x]$ *в* $P[x]$, *т.е. определяет алгебраическую операцию сложения на* $P[x]$.

2. *Формула* fg *для* $f, g \in P[x]$ *задаёт отображение из* $P[x] \times P[x]$ *в* $P[x]$, *т.е. определяет алгебраическую операцию умножения на* $P[x]$.

3. *Формула* αg *для числа* $\alpha \in P$ *(частный случай многочлена) и многочлена* $g \in P[x]$ *задаёт отображение из* $P \times P[x]$ *в* $P[x]$, *т.е. определяет алгебраическую операцию умножения многочлена на число слева.*

Теорема 3.1.1. *Во множестве многочленов* $P[x]$ *с операциями сложения и умножения выполнены свойства:*

1. $\forall f, g, h \in P[x]$ $(f + g) + h = f + (g + h)$ *— ассоциативность сложения,*

2. $\exists 0 \in P[x]\colon \forall f \in P[x]$ $f + 0 = f$ *— существование нулевого элемента,*

3. $\forall f \in P[x]$ $\exists(-f) \in P[x]\colon f + (-f) = 0$ *— существование противоположного элемента,*

4. $\forall f, g \in P[x]$ $f + g = g + f$ *— коммутативность сложения,*

5. $\forall f, g, h \in P[x]$ $(fg)h = f(gh)$ *— ассоциативность умножения,*

6. $\exists 1 \in P[x]\colon \forall f \in P[x]\ f1 = f$ — *существование единичного элемента,*

7. $\forall f, g \in P[x]\ fg = gf$ — *коммутативность умножения,*

8. $\forall f, g, h \in P[x]\ (f+g)h = fh + gh \Leftrightarrow h(f+g) = hf + hg$ — *дистрибутивность умножения относительно сложения.*

9. $\forall f, g \in P[x]$ *из* $fg = 0$ *следует* $f = 0$ *или* $g = 0$ — *отсутствие делителей нуля.*

Следовательно, множество многочленов $P[x]$ *с операциями сложения и умножения является коммутативным кольцом с единицей и без делителей нуля.*

Доказательство.

Докажем некоторые свойства.

Отметим, что нулевым элементом является нулевой многочлен 0, единичным элементом служит многочлен 1, противоположным к многочлену $f = \sum\limits_{k=0}^{\infty} f_k\, x^k$ является многочлен $(-f) = \sum\limits_{k=0}^{\infty} (-f_k)\, x^k$.

5. Докажем ассоциативность умножения $(fg)h = f(gh)$,
где $f = \sum\limits_{k=0}^{\infty} f_k\, x^k,\ g = \sum\limits_{k=0}^{\infty} g_k\, x^k,\ h = \sum\limits_{k=0}^{\infty} h_k\, x^k$.

Обозначим многочлены $u = fg,\ v = uh,\ w = gh,\ y = fw$ с коэффициентами $u_k,\ v_k,\ w_k,\ y_k$ соответственно. Тогда $v = uh = (fg)h$, $y = fw = f(gh)$. Для доказательства равенства $v = y$ покажем, что все соответствующие коэффициенты равны $v_k = y_k$, $k \in Z_+$.

Из формулы (35) определения 4 получим

$$v_k = \sum_{r+s=k} u_r\, h_s = \sum_{r+s=k} \Big(\sum_{j+l=r} f_j\, g_l \Big)\, h_s =$$

$$= \sum_{r+s=k} \Big(\sum_{j+l=r} f_j\, g_l\, h_s \Big) = \sum_{j+l+s=k} f_j\, g_l\, h_s,$$

$$y_k = \sum_{j+r=k} f_j\, w_r = \sum_{j+r=k} f_j \Big(\sum_{l+s=r} g_l\, h_s \Big) = \sum_{j+r=k} \Big(\sum_{l+s=r} f_j\, g_l\, h_s \Big) = \sum_{j+l+s=k} f_j\, g_l\, h_s.$$

Следовательно, $v_k = y_k$ при всех k, тогда $v = y$ и $(fg)h = f(gh)$.

8. Докажем дистрибутивность умножения относительно сложения.

Пусть $f = \sum\limits_{k=0}^{\infty} f_k\, x^k,\ g = \sum\limits_{k=0}^{\infty} g_k\, x^k,\ h = \sum\limits_{k=0}^{\infty} h_k\, x^k$.
Из определения 34 следует:

$$f + g = \sum_{k=0}^{\infty} (f_k + g_k)\, x^k,$$

$$(f+g)h = \sum_{k=0}^{\infty}\left(\sum_{r+s=k}(f_r+g_r)h_s\right)x^k = \sum_{k=0}^{\infty}\left(\sum_{r+s=k}(f_r\,h_s+g_r\,h_s)\right)x^k,$$
$$fh = \sum_{k=0}^{\infty}\left(\sum_{r+s=k}f_r\,h_s\right)x^k,\ \ gh = \sum_{k=0}^{\infty}\left(\sum_{r+s=k}g_r\,h_s\right)x^k,$$
$$fh+gh = \sum_{k=0}^{\infty}\left(\sum_{r+s=k}(f_r\,h_s+g_r\,h_s)\right)x^k.$$

Сравнив коэффициенты многочленов $(f+g)h$ и $fh+gh$, получаем равенство $(f+g)h = fh+gh$.

9. Докажем отсутствие делителей нуля.

Пусть $f, g \in P[x]$ и $fg = 0$. Из равенства (37) получаем $deg(f) + deg(g) = deg(0) = (-\infty)$. Отсюда следует, что $deg(f) = (-\infty)$ или $deg(g) = (-\infty)$, т.е. $f = 0$ или $g = 0$.

Замечания.

1. *Многочлен $f(x) \in P[x]$ имеет обратный $g(x) \in P[x]$ тогда и только тогда, когда $f(x)$ — многочлен нулевой степени.*

Доказательство. Если $f(x)$ — многочлен нулевой степени, т.е. это ненулевое число $f(x) = f_0 \in P$, то возьмём многочлен нулевой степени $g(x) = f_0^{-1}$. Тогда $f(x)g(x) = f_0{f_0}^{-1} = 1$, т.е. обратный многочлен существует.

Если для $f(x)$ существует обратный многочлен $g(x)$, то $f(x)g(x) = 1$. Отсюда следует, что $deg(1) = deg(f(x)) + deg(g(x))$, т.е. $0 = deg(f(x)) + deg(g(x))$. Тогда $deg(f(x)) = deg(g(x)) = 0$, т.е. $f(x)$ — многочлен нулевой степени.

2. *Кольцо $P[x]$ не является полем, т.к. не всякий многочлен имеет обратный многочлен. Поэтому для умножения многочленов обратная операция — деление — не существует.*

3. *Произведение (35) многочленов $f(x) = f_0 + f_1\,x + + \ldots + f_n\,x^n$ степени n и $g(x) = g_0 + g_1\,x + \ldots + g_m\,x^m$ степени m можно найти перемножением двух сумм $(f_0 + f_1\,x + \ldots + f_n\,x^n)(g_0 + g_1\,x + \ldots + g_m\,x^m)$ с раскрытием скобок и последующей группировкой слагаемых с одинаковыми степенями переменной.*

Это следует из дистрибутивности и из того, что многочлены $f(x)$ и $g(x)$ можно рассматривать как суммы одночленов $f_0,\ f_1x,\ \ldots,\ f_nx^n$ и $g_0,\ g_1x,\ \ldots,\ g_mx^m$ соответственно.

Задача 3.1.1.

а) Докажите остальные свойства в теореме 3.1.1.

б) Докажите, что для любых многочленов $f, g, h \in P[x]$ из $f+h = g+h$ следует $f = g$.

в) Докажите, что для любых многочленов $f, g \in P[x]$ *существует, и притом единственное, решение* $y = g + (-f)$ *уравнения* $y + f = g$; *это решение называется* **разностью** g *и* f *и обозначается* $g - f$. *Итак,* $g - f = g + (-f)$ *и эта формула определяет операцию* **вычитания**.

г) Докажите, что для умножение дистрибутивно относительно вычитания, т.е. $\forall f, g, h \in P[x] \quad (f-g)h = fh - gh$ *и* $f(g-h) = fg - fh$.

д) Докажите, что для любых многочленов $f, g, h \in P[x]$: $h \neq 0$ *из* $fh = gh$ *следует* $f = g$.

е) Выясните, когда в неравенстве $deg(f+g) \leqslant \max\{deg(f), deg(g)\}$ *выполняется равенство, и когда — строгое неравенство.*

3.2. Деление многочленов с остатком. Схема Горнера

Для операции умножения многочленов $P[x]$ над полем P обратная операция — деление — не существует. Кольцо $P[x]$ напоминает множество целых чисел Z. Эта аналогия проявляется в том, что для многочленов, как и для целых чисел, имеют место понятия деления нацело, деления с остатком, делителя, наибольшего общего делителя и др.

Приведём обоснование *алгоритма деления с остатком.*

Теорема 3.2.1. *Для любых многочленов* $f(x)$, $g(x) \in P[x]$ $(g(x) \neq \neq 0)$ *существует единственная пара многочленов* $h(x)$, $r(x) \in P[x]$ *такая, что*

$$f(x) = g(x)h(x) + r(x), \tag{38}$$

где либо $r(x) = 0$, *либо* $r(x) \neq 0$ *и* $deg(r) < deg(g)$. (39)

Доказательство. Сначала докажем существование.

При $deg(f) < deg(g)$ напишем тождество $f = g \cdot 0 + f$. Тогда выполнены (38)–(39) при $h = 0$, $r = f$.

Покажем существование при $deg(f) \geqslant deg(g)$. Обозначим $deg(f) = = n$, $deg(g) = m$. Тогда выполнены неравенства $n \geqslant m \geqslant 0$.

Докажем методом математической индукции по степени $n \in Z_+$ утверждение: *для любого* $n \in Z_+$, *для любого многочлена* $f(x)$ *степе-*

ни n и любого многочлена $g(x)$ степени $m\,(0 \leqslant m \leqslant n)$ существуют многочлены $h(x)$ и $r(x)$, для которых выполнены условия (38)–(39).

1) При $n = 0$ имеем $m = 0$, $f(x) = a_0 \neq 0$, $g(x) = b_0 \neq 0$. Отсюда $a_0 = b_0\dfrac{a_0}{b_0} + 0$, т.е. (38)–(39) выполнены при $h(x) = \dfrac{a_0}{b_0}$, $r(x) = 0$.

2) Пусть утверждение верно при всех степенях $n < s$. Докажем утверждение при $n = s$.

Для деления запишем многочлены по убыванию степеней:

$f(x) = a_0x^s + a_1x^{s-1} + \ldots + a_{s-1}x + a_s$ $(s = deg(f))$,

$g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m$ $(m = deg(g),\ s \geqslant m)$.

Применим первый шаг алгоритма деления «уголком» многочлена $f(x)$ на многочлен $g(x)$:

$$\begin{array}{r|l} -\ f(x) & g(x) \\ \cline{2-2} g(x)h_1(x) & h_1(x) \\ \cline{1-1} f_1(x) = f(x) - g(x)h_1(x) & \end{array} \quad \text{или подробнее}$$

$$\begin{array}{r|l} -\ a_0x^s + a_1x^{s-1} + \ldots + a_{s-1}x + a_s & b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m \\ \cline{2-2} a_0x^s + \dfrac{a_0b_1}{b_0}x^{s-1} + \ldots & \dfrac{a_0}{b_0}x^{s-m} \\ \cline{1-1} \left(a_1 - \dfrac{a_0b_1}{b_0}\right)x^{s-1} + \ldots & \end{array}$$

Сначала нашли первое слагаемое частного — одночлен $h_1(x) = \dfrac{a_0}{b_0}x^{s-m}$, потом посчитали произведение $g(x)h_1(x) = \Big(b_0x^m + b_1x^{m-1} + \ldots + b_{m-1}x + b_m\Big)\dfrac{a_0}{b_0}x^{s-m} = a_0x^s + \dfrac{a_0b_1}{b_0}x^{s-1} + \ldots$, затем вычислили разность $f_1(x) = f(x) - g(x)h_1(x) = \Big(a_0x^s + a_1x^{s-1} + \ldots + a_{s-1}x + a_s\Big) - \Big(a_0x^s + \dfrac{a_0b_1}{b_0}x^{s-1} + \ldots\Big) = \Big(a_1 - \dfrac{a_0b_1}{b_0}\Big)x^{s-1} + \ldots$.

Здесь возможны два случая: а) $f_1(x) = 0$, б) $f_1(x) \neq 0$.

а) Из равенства $f_1(x) = f(x) - g(x)h_1(x)$ получим $0 = f(x) - g(x)h_1(x)$, т.е. $f(x) = g(x)h_1(x) + 0$ и тогда выполнены (38)–(39) при $h(x) = h_1(x)$, $r(x) = 0$.

б) Обозначим $s_1 = deg(f_1)$, при этом $0 \leqslant s_1 \leqslant s - 1 < s$. Тогда для многочлена $f_1(x)$ степени меньшей s по предположению индукции найдутся многочлены $h_2(x)$ и $r(x)$, для которых $f_1(x) = g(x)h_2(x) + r(x)$ и

$deg(r(x)) < deg(g(x))$. Отсюда следует $\begin{cases} f_1(x) = f(x) - g(x)h_1(x) \\ f_1(x) = g(x)h_2(x) + r(x) \end{cases} \Rightarrow$ $f(x) - g(x)h_1(x) = g(x)h_2(x) + r(x) \Rightarrow f(x) = g(x)(h_1(x) + h_2(x)) + r(x)$, и выполнены (38)–(39) при $h(x) = h_1(x) + h_2(x)$.

Утверждение о существовании многочленов $h(x)$, $r(x)$, удовлетворяющих условиям (38)–(39) доказано. Теперь докажем их единственность.

Пусть существует ещё одна пара многочленов $h_1(x), r_1(x) \in P[x]$, удовлетворяющих условиям (38)–(39). Тогда $f(x) = g(x)h(x) + r(x)$ и $f(x) = g(x)h_1(x) + r_1(x) \Rightarrow 0 = g(x)(h(x) - h_1(x)) + r(x) - r_1(x) \Rightarrow g(x)(h_1(x) - h(x)) = r(x) - r_1(x)$. Если $r(x) - r_1(x) \neq 0$, то $h_1(x) - h(x) \neq 0$. Рассмотрим в этом случае степени многочленов в равенстве $g(x)(h_1(x) - h(x)) = r(x) - r_1(x)$. В левой части равенства имеем $deg(g(x)(h_1(x) - h(x))) = deg(g(x)) + deg(h_1(x) - h(x)) \geqslant deg(g(x))$. А в правой части равенства в силу $deg(r(x)) < deg(g(x))$ и $deg(r_1(x)) < deg(g(x))$ имеем $deg(r(x) - r_1(x)) < deg(g(x))$, что невозможно. Следовательно, $r(x) - r_1(x) = 0$ и $g(x)(h_1(x) - h(x)) = 0$, а поскольку в кольце $P[x]$ нет делителей нуля, то $h_1(x) - h(x) = 0$. Поэтому $r(x) = r_1(x)$ и $h(x) = h_1(x)$.

Теорема доказана.

Замечание. 1. В практическом применении алгоритма деления «уголком» многочлена $f(x) = a_0x^s + \ldots$ на многочлен $g(x) = b_0x^m + \ldots$ на первом шаге находим многочлен $f_1(x) = f(x) - g(x)h_1(x)$, где $h_1(x) = \dfrac{a_0}{b_0}x^{s-m}$, $deg(f_1) = s_1 < s$. Далее проверяем $s_1 < m$ (сюда входит рассмотренный в доказательстве случай $f_1(x) = 0$) или $s_1 \geqslant m$. Для $s_1 < m$ выразим $f(x) = g(x)h_1(x) + f_1(x)$, тогда выполнены (38)–(39) при $h(x) = h_1(x)$, $r(x) = f_1(x)$, и алгоритм завершен. В противном случае, т.е. для $s_1 \geqslant m$, проводим второй шаг алгоритма (аналогичный первому шагу): делим «уголком» многочлен $f_1(x) = c_0x^{s_1} + \ldots$ на многочлен $g(x)$. Находим $f_2(x) = f(x) - g(x)h_2(x)$, где $h_2(x) = \dfrac{c_0}{b_0}x^{s_1-m}$, $deg(f_2) = s_2 < s_1$. Далее проверяем $s_2 < m$ или $s_2 \geqslant m$ и завершаем алгоритм или продолжаем дальше.

В связи с теоремой 3.2.1 дадим следующее определение.

Определение 1. *Если для многочленов* $f(x), g(x) \in P[x]\,(g(x) \neq \neq 0)$ *существует пара многочленов* $h(x), r(x) \in P[x]$ *такая, что выполнены условия (38)–(39), то многочлены* $h(x)$ *и* $r(x)$ *соответственно называют частным и остатком от деления многочлена* $f(x)$ *на многочлен* $g(x)$.

Кроме общего алгоритма деления «уголком» для нахождения частного и остатка есть *схема Горнера* (*метод Горнера*) для деления многочлена на многочлен первой степени (линейный двучлен).

Пусть заданы многочлены из $P[x]:\ f(x) = a_0x^n + a_1x^{n-1} + \ldots + + a_{n-1}x + a_n$ степени $n \geqslant 1$ и $g(x) = x - c\,(c \in P)$.

Из теоремы 3.2.1 следует, что существует единственная пара многочленов $h(x), r(x) \in P[x]$ такая, что выполнены условия $f(x) = g(x)h(x) + + r(x)$, где либо $r(x) = 0$, либо $r(x) \neq 0$ и $deg(r(x)) < deg(g(x)) = 1$. Отсюда выводим $deg(h(x)) = n - 1$ и либо $r(x) = 0$, либо $r(x) \neq 0$ и $deg(r(x)) = 0$, т. е. $r(x) = r_0$.

Обозначим $h(x) = b_0x^{n-1} + b_1x^{n-2} + \ldots + b_{n-2}x + b_{n-1}$ и $r(x) = r_0$, где b_j, $j = \overline{0, n-1}$ и r_0 — коэффициенты из поля P, которые нужно найти.

Подставим выражения многочленов $g(x)$, $h(x)$, $r(x)$ через коэффициенты в правую часть равенства $f(x) = g(x)h(x) + r(x)$, а затем выражение многочлена $f(x)$ через коэффициенты в левую часть этого равенства:

1) $g(x)h(x) + r(x) = (x - c)(b_0x^{n-1} + b_1x^{n-2} + \ldots + b_{n-2}x + b_{n-1}) + r_0 =$
$= b_0x^n + b_1x^{n-1} + b_2x^{n-2} + \ldots + b_{n-2}x^2 + b_{n-1}x -$
$- cb_0x^{n-1} - cb_1x^{n-2} - \ldots - cb_{n-3}x^2 - cb_{n-2}x - cb_{n-1} + r_0 =$
$= b_0x^n + (b_1 - cb_0)x^{n-1} + (b_2 - cb_1)x^{n-2} + \ldots +$
$+ (b_{n-2} - cb_{n-3})x^2 + (b_{n-1} - cb_{n-2})x + (r_0 - cb_{n-1}),$

2) $f(x) = g(x)h(x) + r(x) \Leftrightarrow$
$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-2}x^2 + a_{n-1}x + a_n =$
$= b_0x^n + (b_1 - b_0c)x^{n-1} + (b_2 - b_1c)x^{n-2} + \ldots +$
$+ (b_{n-2} - b_{n-3}c)x^2 + (b_{n-1} - b_{n-2}c)x + (r_0 - b_{n-1}c).$

Далее приравняем соответствующие коэффициенты равных многочленов в левой и правой частях: $a_0 = b_0$, $a_k = b_k - cb_{k-1}$, $k = \overline{1, n-1}$, $a_n = = r_0 - cb_{n-1}$.

Отсюда получим рекуррентные формулы для последовательного нахождения коэффициентов b_j, $j = \overline{0, n-1}$ и r_0 :

$$\begin{aligned} &b_0 = a_0, \\ &b_k = cb_{k-1} + a_k,\ k = \overline{1, n-1}, \\ &r_0 = cb_{n-1} + a_n. \end{aligned} \tag{40}$$

Здесь $b_0 = a_0$ *и каждый следующий коэффициент* b_k *получается умножением* c *на предыдущий коэффициент* b_{k-1} *и прибавлением коэффициента* a_k. Остаток от деления r_0 получается по этому же закону: $r_0 = cb_{n-1} + a_n$.

При вычислении удобно заполнить таблицу (41). В верхнюю строку таблицы располагаем все коэффициенты a_j, $j = \overline{0, n}$ многочлена $f(x)$. Слева сбоку от нижней строки записываем значение c из многочлена $g(x) = x - c$. В нижнюю строку под коэффициентом a_0 записываем равное ему значение b_0, затем под коэффициентом a_1 записываем сумму $cb_0 + a_1$ и, вычислив его, получаем b_1, и так далее заполняем всю строку.

	a_0	a_1	a_2	$\dots$	a_{n-1}	a_n
c	b_0	$cb_0{+}a_1{=}b_1$	$cb_1{+}a_2{=}b_2$	$\dots$	$cb_{n-2}{+}a_{n-1}{=}b_{n-1}$	$cb_{n-1}{+}a_n{=}r_0$

(41)

Задача 3.2.1.

Докажите, что если для многочленов $f(x), g(x), h(x), r(x) \in P[x]$ *выполнены условия* $f(x) = g(x)h(x) + r(x)$ *и* $deg(r(x)) < deg(g(x))$, *то* $deg(f(x)) = deg(g(x)) + deg(h(x))$.

3.3. Делители многочленов и их свойства

Определение 1. *Если для многочленов* $f(x), g(x) \in P[x]$ *существует многочлен* $h(x) \in P[x]$ *такой, что* $f(x) = g(x)h(x)$, *то говорят, что* $f(x)$ *делится на* $g(x)$ *и* $g(x)$ *делит* $f(x)$, *при этом* $g(x)$ *называют делителем* $f(x)$, *а* $f(x)$ *называют кратным* $g(x)$. *Это записывается следующим образом:* $f(x) \vdots g(x)$ *и* $g(x) \,|\, f(x)$.

Замечания. 1. Если многочлен $f(x)$ делится на многочлен $g(x) \neq \neq 0$, т. е.существует многочлен $h(x)$ такой, что $f(x) = g(x)h(x)$, то многочлен $h(x)$ единственный.

2. Многочлен $f(x) \neq 0$ не делится на 0, т. к. многочлена $h(x)$ такого, что $f(x) = 0h(x)$ не существует.

3. Многочлен 0 делится на 0, т. к. для любого многочлена $h(x)$ верно $0 = 0h(x)$, но результат деления не определён.

Теорема 3.3.1 (Простейшие свойства делителей многочленов).

Для многочленов из $P[x]$ справедливы следующие свойства.

1. *Если $f(x) \vdots g(x)$, а $g(x) \vdots h(x)$, то $f(x) \vdots h(x)$.*
2. *Если $f(x) \vdots \varphi(x)$ и $g(x) \vdots \varphi(x)$, то $(f(x) + g(x)) \vdots \varphi(x)$.*
3. *Если $f(x)$ делится на $\varphi(x)$, то произведение $f(x)$ на любой многочлен $g(x)$ также делится на $\varphi(x)$.*
4. *Если каждый из многочленов $f_j(x)$, $j = \overline{1..k}$ делится на $\varphi(x)$, то на $\varphi(x)$ делится и многочлен $\sum\limits_{j=1}^{k} f_j(x)g_j(x)$, где $g_j(x)$ — произвольные многочлены $j = \overline{1,k}$.*
5. *Всякий многочлен делится на любой многочлен нулевой степени.*
6. *Если $f(x)$ делится на $\varphi(x)$, то $f(x)$ делится и на $c\varphi(x)$, где $c \in$ $\in P$ — любое ненулевое число.*
7. *Многочлены $cf(x)$, где $c \in P$, $c \neq 0$, и только они будут делителями многочлена $f(x)$, имеющими такую же степень, что и $f(x)$.*
8. *Многочлены $f(x)$ и $g(x)$ одновременно делятся друг на друга тогда и только тогда, когда $g(x) = cf(x)$, где $c \in P$, $c \neq 0$.*
9. *Всякий делитель одного из двух многочленов $f(x)$, $cf(x)$, где $c \in$ $P, c \neq 0$, будет являться делителем и для другого многочлена.*
10. *Если $f(x)$ делится на $g(x)$, то $deg(f(x) \geqslant deg(g(x))$.*

Доказательство.

1. $f(x) \vdots g(x) \Leftrightarrow f(x) = g(x)u(x)$ и $g(x) \vdots h(x) \Leftrightarrow g(x) = h(x)v(x)$, где $u(x)$ и $v(x)$ — многочлены. Тогда $f(x) = g(x)u(x) = h(x)v(x)u(x)$, т.е. $f(x) = h(x)(v(x)u(x)) \Rightarrow f(x) \vdots h(x)$.

2. $f(x)\vdots\varphi(x) \Leftrightarrow f(x) = \varphi(x)u(x)$ и $g(x)\vdots\varphi(x) \Leftrightarrow g(x) = \varphi(x)v(x)$, где $u(x)$ и $v(x)$ — многочлены. Тогда $f(x)+g(x) = \varphi(x)u(x)+\varphi(x)v(x)$, т.е. $f(x)+g(x) = \varphi(x)(u(x)+v(x)) \Rightarrow (f(x)+g(x))\vdots\varphi(x)$.

3. $f(x)\vdots\varphi(x) \Leftrightarrow f(x) = \varphi(x)u(x) \Rightarrow f(x)g(x) = \varphi(x)(u(x)g(x)) \Rightarrow$ $\Rightarrow (f(x)g(x))\vdots\varphi(x)$.

4. Докажем методом математической индукции по k.

1) При $k = 1$ $f_1(x)$ делится на $\varphi(x)$. Тогда по свойству 3 $f_1(x)g_1(x)$, делится на $\varphi(x)$.

2) Предположим, что утверждение верно при $k = q$ и докажем, что утверждение верно при $k = q+1$.

Пусть многочлены $f_j(x)$, $j = \overline{1, q+1}$ делятся на $\varphi(x)$. Рассмотрим сумму $\sum\limits_{j=1}^{q+1} f_j(x)g_j(x) = \Big(\sum\limits_{j=1}^{q} f_j(x)g_j(x)\Big) + f_{q+1}(x)g_{q+1}(x)$. На $\varphi(x)$ делится а) первое слагаемое по предположению индукции, б) второе слагаемое по свойству 3, в) их сумма по свойству 2.

5. Если $f(x)$ многочлен и число $c \neq 0$, то $c^{-1}f(x)$ также многочлен и при этом $f(x) = c(c^{-1}f(x)) \Rightarrow f(x)\vdots c$.

6. $f(x)\vdots\varphi(x) \Leftrightarrow f(x) = \varphi(x)u(x) \Rightarrow f(x) = (c\varphi(x))(c^{-1}u(x)) \Rightarrow$ $f(x)\vdots c\varphi(x)$.

7. а) $f(x) = c^{-1}(cf(x)) \Rightarrow f(x)\vdots cf(x)$.

б) Пусть $f(x)\vdots g(x)$ и $deg(f(x)) = deg(g(x))$. Тогда $f(x) = g(x)h(x) \Rightarrow$ $deg(f(x)) = deg(g(x)) + deg(h(x)) = deg(f(x)) + deg(h(x)) \Rightarrow deg(h(x)) =$ $= 0$, т.е. $h(x) = d \in P$, $d \neq 0$. Отсюда $f(x) = g(x)d$ и $g(x) = d^{-1}f(x)$.

8. а) $f(x)\vdots g(x) \Leftrightarrow f(x) = g(x)u(x)$ и $g(x)\vdots f(x) \Leftrightarrow g(x) = f(x)v(x)$. Отсюда следует, что $f = 0 \Leftrightarrow g = 0$ и $f \neq 0 \Leftrightarrow g \neq 0$. В первом случае равенство $g(x) = cf(x)$ верно. Во втором случае из равенств $f(x) =$ $= g(x)u(x)$ и $g(x) = f(x)v(x)$ следует равенство $g(x) = g(x)u(x)v(x)$. Тогда $u(x)v(x) = 1 \Rightarrow deg(u(x)) = deg(v(x)) = 0$. Следовательно, $v(x) =$ $c \in P$, $c \neq 0$ и $g(x) = f(x)v(x) = cf(x)$, т.е. равенство $g(x) = cf(x)$ верно.

б) Если $g(x) = cf(x)$, $c \in P, c \neq 0$, то $f(x) = c^{-1}g(x)$. Тогда $g(x)\vdots f(x)$ и $f(x)\vdots g(x)$.

Задача 3.3.1.

1) Докажите свойства 9–10 в теореме 3.3.1.

2) Пусть $f(x), g(x) \in R[x]$, *где* $f(x) = x^2+2x$, $g(x) = x-b$, $b \in R$.

Докажите, что многочлен $g(x)$ *является делителем многочлена* $f(x)$ *при* $b = 0$ *или* $b = -2$, *а при остальных значениях* $b \in R$ *многочлен* $f(x)$ *не делится на* $g(x)$. *Найдите все делители многочлена* $f(x)$.

3.4. Общий делитель и наибольший общий делитель многочленов. Нахождения наибольшего общего делителя многочленов. Представление наибольшего общего делителя многочленов через эти многочлены

В этом разделе рассматриваются многочлены из $P[x]$ над полем P.

Определение 1. *Пусть заданы многочлены*

$$f_1(x),\ f_2(x),\ \ldots,\ f_k(x). \tag{42}$$

Многочлен $h(x)$ *называется общим делителем многочленов (42), если он является делителем каждого из них, т.е.* $f_j(x) \vdots h(x)$, $j = \overline{1,k}$.

Обозначение: $h(x) \in$ *ОД*$(f_1(x), f_2(x), \ldots, f_k(x))$.

Многочлен $d(x)$ *называется наибольшим общим делителем заданных многочленов (42), если*

1) $d(x)$ *— общий делитель этих многочленов,*

2) $d(x)$ *делится на любой общий делитель* $h(x)$ *этих многочленов.*

Обозначение: $d(x) \in$ *НОД*$(f_1(x), f_2(x), \ldots, f_k(x))$.

Лемма 3.4.1.

1. *Если* $d(x) \in$ *НОД*$(f_1(x), f_2(x), \ldots, f_k(x))$, *то*

   $cd(x) \in$ *НОД*$(f_1(x), f_2(x), \ldots, f_k(x))$ *для любого числа* $c \in P$, $c \neq 0$.

2. *Если* $d_1(x), d_2(x) \in$ *НОД*$(f_1(x), f_2(x), \ldots, f_k(x))$, *то*

   $d_2(x) = cd_1(x)$ *для некоторого числа* $c \in P$, $c \neq 0$.

Доказательство. 1. Пусть $d(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

1) Тогда каждый многочлен $f_j(x) \vdots d(x)$ и из 6-го свойства делителей следует, что каждый многочлен $f_j(x) \vdots cd(x)$ для любого числа $c \in P$, $c \neq$ $\neq 0$, т.е. $cd(x) \in$ ОД$(f_1(x), f_2(x), \ldots, f_k(x))$.

2) Если $h(x)$ общий делитель многочленов (42), то $d(x) \vdots h(x)$. Тогда из 3-го свойства делителей следует, что $cd(x) \vdots h(x)$ для любого числа $c \in P$, $c \neq 0$.

Из 1), 2) получаем $cd(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$ для любого числа $c \in P$, $c \neq 0$.

2. Пусть $d_1(x), d_2(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_k(x))$. Из определения 1 следует, что $d_1(x)$ и $d_2(x)$ являются общими делителями многочленов (42), при этом $d_1(x) \vdots d_2(x)$ и $d_2(x) \vdots d_1(x)$. Тогда в силу 8-го свойства делителей имеем $d_2(x) = cd_1(x)$ для некоторого числа $c \in P$, $c \neq 0$. ■

Следствие.

Если существует многочлен $d(x)$, являющийся наибольшим общим делителем многочленов (42), то все наибольшие общие делители этих многочленов получаются из формулы $cd(x)$, где $c \in P$, $c \neq 0$. ■

Замечание. При $k = 1$ наибольшим общим делителем многочлена $f_1(x)$ является любой многочлен $cf_1(x)$, где $c \in P$, $c \neq 0$. ■

Покажем, что для двух ненулевых многочленов наибольший общий делитель существует и может быть найден с помощью *алгоритма Евклида* или *алгоритма последовательного деления.*

Теорема 3.4.1. *Для любых двух ненулевых многочленов* $f(x), g(x)$ *существует наибольший общий делитель.*

Доказательство. Для построение наибольшего общего делителя многочленов $f(x), g(x)$ выполняем *алгоритм Евклида.*

1-й шаг. Делим f на g : $\begin{array}{r|l} - \; f & g \\ \cline{2-2} \quad gh_1 & h_1 \\ \hline r_1 = f - gh_1 & \end{array}$. Запишем результат деления в виде $f = gh_1 + r_1$, где $deg(r_1) < deg(g)$. Если $r_1 = 0$, то процесс завершён делением нацело на g, причём очевидно, что $g \in$ НОД(f, g). Если $r_1 \neq 0$, то заметим, что степень остатка r_1 меньше степени делителя g, и выполняем шаг 2.

2-й шаг. Делим g на r_1 : $\begin{array}{r|l} - \; g & r_1 \\ \cline{2-2} \quad r_1h_2 & h_2 \\ \hline r_2 = g - r_1h_2 & \end{array}$. Запишем результат деления в виде $g = r_1h_2 + r_2$, где $deg(r_2) < deg(r_1)$. Если $r_2 = 0$, то

процесс завершён делением нацело на r_1. Если $r_2 \neq 0$, то заметим, что степень r_2 меньше степени r_1, и выполняем шаг 3.

3-й шаг. Делим r_1 на r_2 : $\begin{array}{r|l} r_1 & r_2 \\ \cline{2-2} -\ r_2 h_3 & h_3 \\ \hline r_3 = r_1 - r_2 h_3 & \end{array}$. Запишем результат деления в виде $r_1 = r_2 h_3 + r_3$, где $deg(r_3) < deg(r_2)$. Если $r_3 = 0$, то процесс завершён делением нацело на r_2. Если $r_3 \neq 0$, то заметим, что r_3 меньше степени r_2, и выполняем шаг 4, и так далее.

Если процесс завершается на первом шаге, то его результатом является g. В противном случае степени остатков r_j уменьшаются при возрастании j. Поэтому процесс не может продолжаться бесконечно и он завершается делением нацело на некотором $(k+1)$-м шаге: $r_{k-1} = r_k h_{k+1}$. При этом результатом процесса является r_k.

Докажем, что $r_k \in$ НОД(f, g).

Для единообразия обозначим $r_{-1} = f$, $r_0 = g$ и результаты запишем в таблицу.

№ шага	делим	на	получим: частное	получим: остаток	результат деления	условие завершения $r_j = 0$
1	r_{-1}	r_0	h_1	r_1	$r_{-1} = r_0 h_1 + r_1$, $deg(r_1) < deg(r_0)$	если $r_1 \neq 0$, то выполняем шаг 2
2	r_0	r_1	h_2	r_2	$r_0 = r_1 h_2 + r_2$, $deg(r_2) < deg(r_1)$	если $r_2 \neq 0$, то выполняем шаг 3
3	r_1	r_2	h_3	r_3	$r_1 = r_2 h_3 + r_3$, $deg(r_3) < deg(r_2)$	если $r_3 \neq 0$, то выполняем шаг 4
...	...	...	...	...	...	...
$k-1$	r_{k-3}	r_{k-2}	h_{k-1}	r_{k-1}	$r_{k-3} = r_{k-2} h_{k-1} + r_{k-1}$, $deg(r_{k-1}) < deg(r_{k-2})$	если $r_{k-1} \neq 0$, то выполняем шаг k
k	r_{k-2}	r_{k-1}	h_k	r_k	$r_{k-2} = r_{k-1} h_k + r_k$, $deg(r_k) < deg(r_{k-1})$	если $r_k \neq 0$, то выполняем шаг $k+1$
$k+1$	r_{k-1}	r_k	h_{k+1}	$r_{k+1} = 0$	$r_{k-1} = r_k h_{k+1} + r_{k+1} = r_k h_{k+1}$	выполнено

1) Просматривая таблицу снизу вверх, делаем вывод, что каждый многочлен r_{j-2}, $j=\overline{k,1}$ выражается через два последующих многочлена r_{j-1} и r_j : $r_{j-2}=r_{j-1}h_j+r_j$. Из $(k+1)$-й строки следует, что $r_{k-1}=$ $=r_kh_{k+1}$, т.е. $r_{k-1}\vdots r_k$. Тогда из k-й строки следует, что $r_{k-2}\vdots r_k$, из $(k-1)$-й строки следует, что $r_{k-3}\vdots r_k$ и т.д. Получаем r_k — общий делитель всех многочленов r_j, $j=\overline{-1,k}$, в частности, $r_k\in$ ОД$(r_{-1},r_0)=$ ОД(f,g).

2) Просматривая таблицу сверху вниз, получим, что $r_{j+1}=r_{j-1}-$ $-r_jh_{j+1}$ для $j=\overline{0,k-1}$, т.е. каждый многочлен r_{j+1} выражается через два предыдущих многочлена r_j и r_{j-1}. Поэтому любой общий делитель первых двух многочленов $r_{-1}=f$ и $r_0=g$ является делителем всех многочленов r_j для $j=\overline{1,k}$. В частности, r_k делится на любой общий делитель многочленов f и g.

Итак, $r_k\in$ НОД(f,g) и теорема доказана. ■

С помощью алгоритма Евклида докажем утверждение о формуле представления наибольшего общего делителя двух многочленов через эти два многочлена.

Теорема 3.4.2. *Если $d(x)$ — наибольший общий делитель многочленов $f(x)$ и $g(x)$, то существуют такие многочлены $u(x)$ и $v(x)$, что*

$$d(x)=f(x)u(x)+g(x)v(x), \tag{43}$$

Если степени многочленов $f(x)$ и $g(x)$ больше нуля, то можно считать при этом, что $deg(u(x))<deg(g(x))$ и $deg(v(x))<deg(f(x))$.

Доказательство. Если f делится на g, т.е. $f=gh_1$, где h_1 — многочлен, то $d=g\in$ НОД(f,g), и формула (43) выполнена при $u=0$, $v=1$.

Если f не делится на g, то проделаем $(k+1)$ шагов алгоритма Евклида и получим $r_k\in$ НОД(f,g).

Выпишем формулы вычисления ненулевых остатков r_j :

$r_1=f-gh_1$,

$r_2=g-r_1h_2$, $r_2=g-(f-gh_1)h_2=(-h_2)f+(1+h_1h_2)g$,

далее каждый следующий остаток выражается через два предыдущих

$r_{j+1} = r_{j-1} - r_j h_{j+1}$ для $j = \overline{0, k-1}$.

С помощью этих формул будем в порядке возрастания номеров j от 1 до k последовательно выражать через f и g остатки $r_j = fu_j + gv_j$. При $j = k$ получим $r_k = fu_k + gv_k$, т.е. равенство (43) будет выполнено для $d = r_k$ при $u = u_k$, $v = v_k$.

В силу леммы 3.4.1 любой наибольший общий делитель d многочленов f и g выражается через r_k : $d = c \cdot r_k$, $c = const \neq 0$. Тогда $c \cdot r_k = f \cdot (c \cdot u_k) + g \cdot (c \cdot v_k)$, т.е. $d = f \cdot u + g \cdot v$ при $u = c \cdot u_k$, $v = c \cdot v_k$.

Для доказательства второго утверждения теоремы предположим, что равенство (43) выполнено, но, например, $deg(u(x)) \geqslant deg(g(x))$. Разделим $u(x)$ на $g(x)$:

$$u(x) = g(x)w(x) + u_1(x),$$

где $deg(u_1(x)) < deg(g(x))$, и подставим это выражение в (43).

Получим $d(x) = f(x)g(x)w(x) + f(x)u_1(x) + g(x)v(x)$

или $d(x) = f(x)u_1(x) + g(x)(f(x)w(x) + v(x))$.

Отсюда следует другое представление для $d(x)$:

$$d(x) = f(x)u_1(x) + g(x)v_1(x), \quad \text{где} \quad v_1(x) = f(x)w(x) + v(x). \qquad (44)$$

Здесь уже выполнено одно требуемое неравенство $deg(u_1(x)) < deg(g(x))$. Докажем второе неравенство $deg(v_1(x)) < deg(f(x))$ методом «от противного».

Предположим противное: $deg(v_1(x)) \geqslant deg(f(x))$. Покажем, что из этого предположения следует вывод: $deg(d) \geqslant deg(f) + deg(g)$, приводящий к противоречию.

Оценим в (44) степень произведения многочленов $g(x)v_1(x)$:

$$deg(g(x)v_1(x)) = deg(g(x)) + deg(v_1(x)) \geqslant deg((g(x)) + deg(f(x)).$$

Оценим в (44) степень произведения многочленов $f(x)u_1(x)$:

$$deg(f(x)u_1(x)) = deg((f(x)) + deg(u_1(x)) < deg(f(x)) + deg(g(x)).$$

Из двух неравенств получаем $deg(g(x)v_1(x)) > deg(f(x)u_1(x))$.

Очевидно, что степень суммы двух многочленов разных степеней равна их максимальной степени. Следовательно, $deg(f(x)u_1(x) + g(x)v_1(x)) = deg(g(x)v_1(x))$. Тогда из (44) получаем $deg(d(x)) =$

$= deg(f(x)u_1(x) + g(x)v_1(x)) = deg(g(x)v_1(x)) \geqslant deg(g(x)) + deg(f(x))$, т.е. $deg(d(x)) \geqslant deg(g(x)) + deg(f(x))$.

По условию $d(x) \in$ НОД$(f(x), g(x))$, поэтому $f(x)$ делится на $d(x)$. Следовательно, $deg(f(x)) \geqslant deg(d(x))$ и, учитывая неравенство $deg(d(x)) \geqslant deg(g(x)) + deg(f(x))$, получаем $deg(f(x)) \geqslant deg(g(x)) + deg(f(x))$. Отсюда $deg(g(x)) = 0$, что противоречит $deg(g(x)) > 0$.

Полученное противоречие доказывает, что для многочлена $v_1(x)$ в представлении (44) выполнено неравенство $deg(v_1(x)) < deg(f(x))$. И ранее было получено неравенство $deg(u_1(x)) < deg(g(x))$.

Теорема доказана. ■

Существование наибольшего общего делителя для системы многочленов (42) при произвольном $k \geqslant 3$ и способ его вычисления получается из следующей теоремы.

Теорема 3.4.3.

Для любых ненулевых многочленов $f_1(x), f_2(x), \ldots, f_k(x)$ при произвольном $k \in N$, $k \geqslant 2$ справедливо равенство

$$НОД(f_1(x), f_2(x), \ldots, f_k(x)) = НОД(НОД(f_1(x), \ldots, f_{k-1}(x)), f_k(x)). \quad (45)$$

Это означает, что

- *существует многочлен $d_{k-1}(x)$, который является наибольшим общим делителем многочленов $f_1(x), f_2(x), \ldots, f_{k-1}(x)$,*
- *существует многочлен $d_k(x)$, который является наибольшим общим делителем многочленов $d_{k-1}(x), f_k(x)$,*
- *многочлен $d_k(x)$, является наибольшим общим делителем многочленов $f_1(x), f_2(x), \ldots, f_k(x)$ и выполнено равенство (45).*

Согласно этой теореме для нахождения НОД$(f_1(x), f_2(x), \ldots, f_k(x))$ сначала берём $d_1(x) = f_1(x)$, далее находим $d_2(x) \in$ НОД$(d_1(x), f_2(x))$, после этого определяем $d_j(x) \in$ НОД$(d_{j-1}(x), f_j(x))$, $j = \overline{3, k}$. Ответом является множество многочленов $cd_k(x)$, где $c = const \neq 0$.

Доказательство теоремы 3.4.3 проводится методом математической индукции.

При $k = 2$ утверждение следует из теоремы 3.4.2.

Далее доказываем, что из предположения о справедливости утверждения при $k = m$ следует его справедливость при $k = m + 1$.

По предположению индукции существует многочлен $d_m(x)$, который является наибольшим общим делителем многочленов $f_1(x)$, $f_2(x)$, $\ldots, f_m(x)$. По теореме 3.4.1 существует многочлен $d_{m+1}(x)$, который является наибольшим общим делителем многочленов $d_m(x), f_{m+1}(x)$. Покажем, что $d_{m+1}(x)$ является наибольшим общим делителем многочленов $f_1(x)$, $f_2(x)$, $\ldots$, $f_{m+1}(x)$.

Во-первых, все $f_j(x)$, $j = \overline{1, m}$ делятся на $d_m(x)$, а $d_m(x)$ и $f_{m+1}(x)$ делятся на $d_{m+1}(x)$, поэтому все $f_j(x)$, $j = \overline{1, m+1}$ делятся на $d_{m+1}(x)$. Во-вторых, любой общий делитель $h(x)$ многочленов $f_1(x)$, $f_2(x)$, $\ldots$, $f_{m+1}(x)$ является делителем $d_m(x)$ и $f_{m+1}(x)$, и поэтому является делителем $d_{m+1}(x)$. Следовательно, $d_{m+1}(x)$ является наибольшим общим делителем всех многочленов $f_j(x)$, $j = \overline{1, m+1}$.

В силу леммы 3.4.1,
$\forall c_1 \in P$, $c_1 \neq 0$ выполнено НОД$(c_1 d_m(x), f_{m+1}(x)) = \{c_2 d_{m+1}(x) |\, c_2 \in \in P,\, c_2 \neq 0\}$ и НОД$(f_1(x), f_2(x), \ldots, f_{m+1}(x)) = \{c_3 d_{m+1}(x) |\, c_3 \in P,\, c_3 \neq \neq 0\}$, т. е. при $k = m + 1$ справедливо равенство (45). ■

Справедлив аналог теоремы 3.4.2 о формуле представления наибольшего общего делителя нескольких многочленов через эти многочлены.

Теорема 3.4.4. *Если $d(x)$ — наибольший общий делитель многочленов $f_1(x), f_2(x), \ldots, f_k(x)$ при произвольном $k \in \mathbb{N}$, $k \geqslant 2$, то существуют такие многочлены $u_1(x), u_2(x), \ldots, u_k(x)$, что*

$$d(x) = f_1(x)u_1(x) + f_2(x)u_2(x) + \ldots + f_k(x)u_k(x). \qquad (46)$$

Доказательство теоремы 3.4.4 проводится методом математической индукции.

При $k = 2$ утверждение следует из теоремы 3.4.2.

Далее докажем, что из предположения о справедливости утверждения для $D_m(x) \in$ НОД$(f_1(x), f_2(x), \ldots, f_m(x))$ следует его справедливость для $D_{m+1}(x) \in$ НОД$(D_m(x), f_{m+1}(x)) =$ НОД$(f_1(x)$, $f_2(x)$, $\ldots$, $f_{m+1}(x))$ ($D_{m+1}(x)$ строится как в теореме 3.4.3).

По предположению индукции для многочлена $D_m(x)$ существуют многочлены $w_j(x)$, $j = \overline{1,m}$ такие, что $D_m(x) = \sum\limits_{j=1}^{m} f_j(x)w_j(x)$. По теореме 3.4.2 существуют такие многочлены $u(x)$ и $v(x)$, что $D_{m+1}(x) = D_m(x)u(x) + f_{m+1}(x)v(x)$. Тогда $D_{m+1}(x) = \sum\limits_{j=1}^{m} f_j(x)(w_j(x)u(x)) + f_{m+1}(x)v(x)$. Обозначив $u_j(x) = w_j(x)u(x)$, $j = \overline{1,m}$ и $u_{m+1} = v(x)$, получим для $k = m+1$ и $D_{m+1}(x)$ представление (46): $D_{m+1}(x) = \sum\limits_{j=1}^{m+1} f_j(x)u_j(x)$. Поскольку любой наибольший общий делитель $d(x)$ многочленов $f_1(x), f_2(x), \ldots, f_{m+1}(x)$ равен $cD_{m+1}(x)$ для некоторого числа $c \in P$, $c \neq 0$, тогда для $d(x)$ справедливо представление $d(x) = \sum\limits_{j=1}^{m+1} f_j(x)(cu_j(x))$. Теорема доказана. ■

Задача 3.4.1.

1) Докажите свойства 9–10 в теореме 3.3.1.

2) Пусть $f(x), g(x) \in \mathbb{R}[x]$, *причём* $f(x) = x^2+2x$, $g(x) = x-b$, $b \in \mathbb{R}$.

Докажите, что многочлен $g(x)$ *является делителем многочлена* $f(x)$ *при* $b = 0$ *или* $b = -2$, *а при остальных значениях* $b \in \mathbb{R}$ *многочлен* $f(x)$ *не делится на* $g(x)$. *Найдите все делители многочлена* $f(x)$.

3.5. Взаимная простота многочленов. Свойства, связанные со взаимной простотой

В этом разделе рассматриваются многочлены из $P[x]$ над полем P.

Каждый многочлен делится на любой многочлен нулевой степени

$$d(x) = c,\ c \in P,\ c \neq 0. \tag{47}$$

Поэтому любой многочлен нулевой степени является общим делителем произвольных многочленов.

Определение 1. *Пусть заданы многочлены*

$$f_1(x),\ f_2(x),\ \ldots,\ f_k(x). \tag{48}$$

Многочлены (48) называются взаимно простыми, если их общими делителями являются только многочлены нулевой степени.

Пример 1. Многочлены $f_1(x) = x^2 - 4x$ и $f_2(x) = x^2 - 4$ делятся на $x - 4$, и поэтому не являются взаимно простыми.

Пример 2. Многочлен $f_1(x) = x^2 + 2x$ делится только на $c_1 \cdot 1$, $c_2 x$, $c_3(x+2)$, $c_4(x^2+2x)$, многочлен $f_2(x) = x^2 - 2x + 1$ делятся только на $c_5 \cdot 1$, $c_6(x-1)$, $c_7(x-1)^2$, и поэтому $f_1(x)$ и $f_2(x)$ являются взаимно простыми.

Пример 3. Возьмём многочлены $f_1(x) = x(x+1)$, $f_2(x) = (x+1)(x+2)$, $f_3(x) = x(x+2)$. Любые два из них не являются взаимно простыми. Действительно, $f_1(x), f_2(x)$ делятся на $x+1$, $f_2(x), f_3(x)$ делятся на $x+2$, $f_1(x), f_3(x)$ делятся на x. Многочлены $f_1(x), f_2(x), f_3(x)$ являются взаимно простыми (докажите!).

Теорема 3.5.1. *(Критерии взаимной простоты)*

Следующие условия эквивалентны:

(а) *многочлены (48) взаимно просты;*

(б) *многочлен* 1 *является наибольшим общим делителем многочленов (48);*

(в) *существуют такие многочлены* $u_j(x)$, $j = \overline{1,k}$, *что*

$$\sum_{j=1}^{k} f_j(x) u_j(x) = 1.$$

Доказательство.

(а)$\Rightarrow$(б). Пусть выполнено (а), тогда все общие делители многочленов (48) имеют вид (47). Поэтому наибольшим общим делителем многочленов (48) является любой многочлен $d(x) = c$, $c \in P$, $c \neq 0$, в частности, и при $c = 1$.

(б)$\Rightarrow$(в). Если выполнено (б), то из теоремы 3.4.3 следует (в).

(в)$\Rightarrow$(а). Пусть выполнено (в) и многочлен $d(x)$ является общим делителем многочленов (48). Тогда $(\sum\limits_{j=1}^{k} f_j(x)u_j(x)) \vdots d(x) \Rightarrow 1 \vdots d(x) \Rightarrow$ $\Rightarrow deg(1) \geqslant deg(d(x)) \Rightarrow deg(d(x)) = 0 \Rightarrow d(x) = c$, $c \in P$, $c \neq 0 \Rightarrow$ (а).

Итак, (а) $\Leftrightarrow$ (б) $\Leftrightarrow$ (в). ■

Теорема 3.5.2. *(Свойства, связанные со взаимной простотой)*

1. *Если многочлен* $f(x)$ *взаимно прост с каждым из многочленов* $g_1(x)$ *и* $g_2(x)$, *то* $f(x)$ *взаимно прост и с их произведением* $g_1(x)g_2(x)$.

2. *Если произведение многочленов* $f_1(x)f_2(x)$ *делится на многочлен* $g(x)$, *но* $f_1(x)$ *и* $g(x)$ *взаимно просты, то* $f_2(x)$ *делится на* $g(x)$.

3. *Если многочлен* $f(x)$ *делится на каждый из многочленов* $g_1(x)$ *и* $g_2(x)$, *которые между собой взаимно просты, то* $f(x)$ *делится и на их произведение* $g_1(x)g_2(x)$.

Доказательство.

1. Многочлены $f(x)$ и $g_1(x)$ взаимно просты $\Rightarrow \exists u(x), v(x) \in P[x]$ такие, что $f(x)u(x) + g_1(x)v(x) = 1$. Умножим обе части равенства на $g_2(x)$, и получим $f(x)(u(x)g_2(x)) + (g_1(x)g_2(x))v(x) = g_2(x)$. Отсюда следует, что любой общий делитель $d(x)$ многочленов $f(x)$ и $g_1(x)g_2(x)$ является делителем $g_2(x)$. Тогда $d(x)$ является общим делителем многочленов $f(x)$ и $g_2(x)$, а они взаимно просты по условию, и поэтому $d(x) = c$, $c \in P$, $c \neq 0$. Итак, только многочлены нулевой степени являются общими делителями многочленов $f(x)$ и $g_1(x)g_2(x)$, т.е. они взаимно просты.

2. Многочлены $f_1(x)$ и $g(x)$ взаимно просты $\Rightarrow \exists u(x), v(x) \in P[x]$ такие, что $f_1(x)u(x) + g(x)v(x) = 1$. Умножим обе части равенства на $f_2(x)$, и получим $(f_1(x)f_2(x))u(x) + g(x)(v(x)f_2(x)) = f_2(x)$. Оба слагаемых левой части делятся на $g(x)$, а тогда на него делится и $f_2(x)$.

3. Многочлен $f(x)$ делится на многочлен $g_1(x)$, поэтому верно равенство $f(x) = g_1(x)h_1(x)$. Многочлен $f(x)$ делится на многочлен $g_2(x)$, следовательно, произведение $g_1(x)h_1(x)$ делится на $g_2(x)$, при этом $g_1(x)$ и $g_2(x)$ между собой взаимно просты. Тогда из свойства 2 следует, что $h_1(x)$ делится на $g_2(x)$, т.е. $h_1(x) = g_2(x)h_2(x)$. Поэтому верно равенство $f(x) = g_1(x)h_1(x) = g_1(x)g_2(x)h_2(x)$, откуда следует, что $f(x)$ делится на произведение $g_1(x)g_2(x)$. ■

Следствия. а) *Если многочлены* $f(x)$ *и* $g(x)$ *взаимно просты, то взаимно просты многочлен* $f(x)$ *и любая степень многочлена* $g^n(x)$, *где* $n \in N$.

б) *Если многочлены* $f(x)$ *и* $g(x)$ *взаимно просты, то взаимно просты любые их степени* $f^m(x)$ *и* $g^n(x)$, *где* $m, n \in N$.

Доказательство. а) Докажем методом математической индукции по n.

1) При $n = 1$ утверждение верно.

2) Докажем, что из предположения о справедливости утверждения для $n = k$, следует его справедливость при $n = k + 1$.

Действительно, $f(x)$ и $g(x)$ взаимно просты по условию, $f(x)$ и $g^k(x)$ взаимно просты по предположению индукции. Тогда из свойства 1 следует, что $f(x)$ взаимно прост с произведением $g^k(x)$ и $g(x)$, т.е. $f(x)$ взаимно прост с $g^{k+1}(x)$.

б) Докажем методом математической индукции по m следующее утверждение: *для каждого $m \in N$ степень $f^m(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in N$.*

1) При $m = 1$ утверждение верно в силу следствия 1а.

2) Пусть утверждение верно при $m = k$, т.е. степень $f^k(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in N$.

Представим степень $f^{k+1}(x)$ в виде произведения $f^{k+1}(x) = f^k(x) \times \times f(x)$. По предположению индукции степень $f^k(x)$ взаимно проста со степенью $g^n(x)$ при всех $n \in N$ и из а) $f(x)$ взаимно прост со степенью $g^n(x)$ при всех $n \in N$. Тогда отсюда и из свойства 1 следует, что произведение $f^k(x)f(x)$ взаимно просто с $g^n(x)$ при всех $n \in N$, т.е. утверждение верно при $m = k + 1$. ■

Задача 3.5.1. *1) Докажите, что многочлены $f_1(x)$, $f_2(x)$, $f_3(x)$ из примера 3 взаимно просты.*

2) Докажите, что если любые два многочлена из (48) взаимно просты, то все они взаимно просты.

3) Докажите обобщение свойства 1 теоремы 3.5.2 на любое число многочленов $g_j(x), k = \overline{1, k}$.

4) Докажите, что если взаимно просты некоторые степени $f^{m_0}(x)$ и $g^{n_0}(x)$ многочленов $f(x)$ и $g(x)$, то взаимно просты любые их степени $f^m(x)$ и $g^n(x)$, где $m, n \in \mathbb{N}$.

5) Пусть $d(x) = f(x)u(x) + g(x)v(x)$, где $d(x)$ — наибольший общий делитель многочленов $f(x)$ и $g(x)$. Найдите наибольший общий делитель многочленов $u(x)$ и $v(x)$. Ответ обоснуйте.

3.6. Корни многочленов. Кратность корня

В разделе 3.1. мы определили многочлен $f(x) \in P[x]$ над полем P как некоторое формальное выражение, которое вполне определяется набором своих коэффициентов. Однако многочлен можно рассматривать и как функцию от переменной x.

Определение 1. *Если* $f(x) = \sum\limits_{k=0}^{n} f_k x^k$ *— многочлен над полем* P, c *— некоторое число из поля* P, *то число* $f(c) = \sum\limits_{k=0}^{n} f_k\, c^k$ *называется значением многочлена* $f(x)$ *при* $x = c$.

Корнем многочлена $f(x) \in P[x]$ *или уравнения* $f(x) = 0$ *называется такое число* $c \in P$, *что* $f(c) = 0$.

Из теории функций известно, что две функции называются равными, если их значения равны при всех значениях переменной. Очевидно что если многочлены $f(x)$ и $g(x)$ равны как многочлены, т.е. их коэффициенты совпадают (см. определение 2 стр. 61), то они равны как функции. Обратное утверждение требует дополнительного исследования и это будет сделано позже для многочленов $C[x]$. А пока, забегая вперёд, будем иметь в виду, что оба подхода к понятию равенства многочленов совпадают.

Сказанное относится и к операциям над многочленами. Если определены (см. определение 4 стр. 62) сумма и произведение многочленов $f(x)$ и $g(x)$: $\varphi(x) = f(x) + g(x)$, $\psi(x) = f(x)g(x)$, то $\forall\, c \in P$ справедливы равенства $\varphi(c) = f(c) + g(c)$, $\psi(c) = f(c)g(c)$, которые соответствуют определению суммы и произведения функций. Следовательно, операциям сложения и умножения над многочленами соответствуют операции сложения и умножения над функциями, которые определяются этими многочленами.

Теорема 3.6.1 (Теорема Безу).
Остаток от деления многочлена $f(x)$ *на* $x - c$ $(c \in P)$ *равен* $f(c)$.

Доказательство. Разделим многочлен $f(x)$ на $x - c$. Тогда $f(x) = (x - c)h(x) + r(x)$, где $\deg(r(x)) < \deg(x - c) = 1$, так что $r(x) = r_0$ —

константа. Из $f(x) = (x-c)h(x) + r_0$ получим при $x = c$ значение $f(c) = (c-c)h(c) + r_0 = r_0$, т.е. остаток $r_0 = f(c)$.

Следствия. 1. *Число $c \in P$ является корнем многочлена $f(x) \in$ $\in P[x]$ тогда и только тогда, когда многочлен $f(x)$ делится на $x - c$, т. е. $f(x) = (x-c)h(x)$, где многочлен $h(x) \in P[x]$.*

2. *Если число $c \in P$ является корнем ненулевого многочлена $f(x)$ из $P[x]$, то существуют число $k \in N$ и многочлен $s(x) \in P[x]$ такие, что*

$$f(x) = (x-c)^k s(x), \tag{49}$$

где многочлен $s(x)$ не делится на $x - c$, т. е. число c своим корнем не имеет.

Доказательство.

1. Из теоремы Безу следует, что $f(x) = (x-c)h(x) + f(c)$.

Если $f(c) = 0$, то $f(x) = (x-c)h(x)$ и $f(x)$ делится на $x - c$.

Если $f(x)$ делится на $x - c$, то $f(x) = (x-c)h(x)$. Отсюда при $x = c$ получим $f(c) = 0$.

2. Докажем следствие 2 методом математической индукции по степени m многочлена $f(x)$.

Пусть многочлен $f(x) = a_0x + a_1$ $(a_0 \neq 0)$ степени $m = 1$ имеет корень $x = c$. Тогда $a_0c + a_1 = 0 \Rightarrow a_1 = -a_0c$ и $f(x) = a_0x - a_0c =$ $= (x-c)a_0$. Поэтому равенство $f(x) = (x-c)^k s(x)$ выполнено при $k =$ $= 1$, $s(x) = a_0 \neq 0$.

Покажем, что из справедливости утверждения для степени $m = n$ следует его справедливость для степени $m = n + 1$.

Из следствия 1 имеем $f(x) = (x-c)h(x)$, где $h(x)$ — многочлен из $P[x]$. Приравняем степени $\deg(f(x)) = \deg(x-c) + \deg(h(x)) \Rightarrow n + 1 =$ $= 1 + \deg(h(x)) \Rightarrow \deg(h(x)) = n$. Если $h(c) \neq 0$, то утверждение верно при $k = 1$, $s(x) = h(x)$. Если $h(c) = 0$, то по предположению индукции существуют число $k_1 \in N$ и многочлен $h_1(x) \in P[x]$ такие, что $h(x) = (x-c)^{k_1}h_1(x)$, причём $h_1(c) \neq 0$. Тогда $f(x) = (x-c)h(x) =$ $= (x-c)(x-c)^{k_1}h_1(x) = (x-c)^{1+k_1}h_1(x)$, где $h_1(c) \neq 0$. Поэтому утверждение верно при $k = 1 + k_1$, $s(x) = h_1(x)$.■

Замечание. Следствие 1 из теоремы Безу говорит о том, что задача о нахождении корня $x = c$ многочлена эквивалентна задаче о нахождении делителя $x - c$ этого многочлена. ■

Определение 2. *Число $k \in N$ называется кратностью корня $c \in P$ многочлена $f(x) \in P[x]$, если выполнено равенство*

$$f(x) = (x - c)^k s(x),$$

где многочлен $s(x) \in P[x]$ не делится на $x - c$, т.е. число c своим корнем не имеет. При этом сам корень c называется k-кратным. Для $k = 1$ говорят, что корень c — простой. ■

Если для многочленов над полем P определено понятие производных, справедливы формулы дифференцирования суммы, произведения и формула Тейлора, то тогда справедливы следующие утверждения о кратности корня многочлена.

Теорема 3.6.2.

1° *Если число c является k-кратным корнем многочлена $f(x)$, то при $k > 1$ оно будет $(k-1)$-кратным корнем первой производной этого многочлена; при $k = 1$ число c не является корнем $f'(x)$.*

2° *Если число c является k-кратным корнем многочлена $f(x)$, то c будет $(k-m)$-кратным корнем m-й производной этого многочлена при $m < k$ и не будет корнем k-й производной от $f(x)$.*

3° *Число c является k-кратным корнем многочлена $f(x)$ тогда и только тогда, когда $f(c) = f'(c) = \cdots = f^{(k-1)}(c) = 0$ и $f^{(k)}(c) \neq 0$.*

Доказательство. 1° Пусть c является k-кратным корнем многочлена $f(x)$. Тогда $f(x) = (x - c)^k s(x)$, где многочлен $s(x) \in P[x]$ не делится на $x - c$, т.е. $s(c) \neq 0$. Продифференцируем $f'(x) = k(x - c)^{k-1} s(x) + (x - c)^k s'(x) = (x - c)^{k-1}\big(ks(x) + (x - c)s'(x)\big)$, обозначим $h(x) = ks(x) + (x - c)s'(x)$ и получим $f'(x) = (x - c)^{k-1} h(x)$. Для многочлена $h(x)$ имеем $h(c) = ks(c) \neq 0$. Поэтому $x = c$ будет $(k-1)$-кратным корнем $f'(x)$ при $k > 1$ и число c не является корнем $f'(x)$ при $k = 1$.

2° Применим метод математической индукции по k.

1) Утверждение верно при $k = 1$ в силу 1°.

2) Покажем, что из справедливости утверждения при всех $k < p$ следует справедливость утверждения при $k = p$.

Если число c является p-кратным корнем многочлена $f(x)$, то в силу 1° число c будет $(p-1)$-кратным корнем $g(x) = f'(x)$ при $p > 1$. Из предположения индукции число c будет $((p-1)-(m-1))$-кратным корнем $g^{(m-1)}(x)$ при $m-1 < p-1$ и не будет корнем $g^{(p-1)}(x)$. Поэтому число c будет $(p-m)$-кратным корнем $f^{(m)}(x)$ при $m < p$ и не будет корнем $f^{(p)}(x)$.

3° Если число c является k-кратным корнем многочлена $f(x)$, то в силу 2° $f(c) = f'(c) = \ldots = f^{(k-1)}(c) = 0$ и $f^{(k)}(c) \neq 0$.

Обратно, пусть для многочлена степени n выполнены равенства $f(c) = f'(c) = \ldots = f^{(k-1)}(c) = 0$ и $f^{(k)}(c) \neq 0$. Применим формулу Тейлора $f(x) = \sum\limits_{m=0}^{n} \dfrac{f^{(m)}(c)}{m!}(x-c)^m$ (по условию $k \leqslant n$).

Поскольку $f(c) = f'(c) = \ldots = f^{(k-1)}(c) = 0$ и $f^{(k)}(c) \neq 0$, имеем

$$f(x) = \frac{f^{(k)}(c)}{k!}(x-c)^k + \frac{f^{(k+1)}(c)}{(k+1)!}(x-c)^{k+1} + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n =$$

$$= (x-c)^k\Big(\frac{f^{(k)}(c)}{k!} + \frac{f^{(k+1)}(c)}{(k+1)!}(x-c)^1 + \ldots \frac{f^{(n)}(c)}{n!}(x-c)^{n-k}\Big) = (x-c)^k s(x),$$

где $s(c) = \dfrac{f^{(k)}(c)}{k!} \neq 0$. Отсюда $x = c$ является k-кратным корнем многочлена $f(x)$. ■

3.7. Основная теорема алгебры для многочленов над C

Теорема 3.7.1. *Любой многочлен $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ с комплексными коэффициентами степени $n \geqslant 1$ имеет хотя бы один корень, в общем случае комплексный.*

Доказательство теоремы. Покажем, что инфимум значений модуля многочлена $|f(z)|$ достигается при некотором $z = z_0$ и при этом $|f(z_0)| = 0$, т. е. z_0 — корень многочлена $f(z)$.

Для этого нам понадобятся дополнительные понятия и факты.

Сначала докажем непрерывность многочлена и его модуля.

Комплексная функция $f(z)$ от комплексной переменной z называется непрерывной в точке z_0, если малому изменению аргумента соответствует малое изменение функции, т.е. для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех z, для которых $|z - z_0| < \delta$, выполняется неравенство $|f(z) - f(z_0)| < \varepsilon$.

Лемма 3.7.1. (**Непрерывность в нуле многочлена с нулевым свободным членом.**)

Пусть $f(z) = a_1 z + a_2 z^2 + \ldots + a_n z^n$ — многочлен с нулевым свободным членом. Тогда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех z, для которых $|z| < \delta$, выполняется неравенство $|f(z)| < \varepsilon$.

Доказательство леммы 3.7.1. Пусть $|z| < 1$, тогда $|f(z)| = |z||a_1 + a_2 z^2 + \ldots + a_n z^{n-1}| \leqslant |z|(|a_1| + |a_2||z|^2| + \ldots + |a_n||z|^{n-1}) \leqslant |z|(|a_1| + |a_2| + \ldots + |a_n|)$. Правая часть мала при малых z. Положим $M = |a_1| + |a_2| + \ldots + |a_n|$ и возьмём $\delta = \min(1, \dfrac{\varepsilon}{M})$. Тогда при $|z| < \delta$ имеем $|f(z)| \leqslant |z|M < \dfrac{\varepsilon M}{M} = \varepsilon$. ■

Лемма 3.7.2. (**Непрерывность многочлена.**)

Многочлен $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ есть непрерывная функция во всех точках комплексной плоскости.

Доказательство леммы 3.7.2. Пусть z_0 — произвольное комплексное число. Разложим $f(z)$ по степеням $z - z_0$: $f(z) = c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n$. Тогда $c_0 = f(z_0)$, так что $f(z) - f(z_0) = c_1(z - z_0) + \ldots + c_n(z - z_0)^n$.

Правая часть представляет собой многочлен от $z - z_0$ с нулевым свободным членом. По лемме 3.7.1 правая часть мала при малых $z - z_0$. Действительно, для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f(z) - f(z_0)| < \varepsilon$ для всех z, для которых $|z - z_0| < \delta$. ■

Лемма 3.7.3. (**Непрерывность модуля многочлена.**)

Модуль многочлена есть непрерывная функция.

Доказательство леммы 3.7.3. Утверждение вытекает из леммы 3.7.2 и неравенства $\big||f(z)| - |f(z_0)|\big| \leqslant |f(z) - f(z_0)|$, т.к. правая часть мала при малых $z - z_0$.■

Лемма 3.7.4. (**Значения многочлена $f(z) \to \infty$ при $z \to \infty$.**)
Если $f(z) = a_0 + a_1 z + \ldots + a_n z^n$ многочлен степени $n \geqslant 1$, то для любого $M > 0$ существует $R > 0$ такое что для всех z, для которых $|z| > R$, выполняется неравенство $|f(z)| > M$.

Доказательство леммы 3.7.4. Преобразуем $f(z) = a_0 + a_1 z + \ldots + a_n z^n = a_n z^n (1 + \frac{a_{n-1}}{a_n} z^{-1} + \ldots + \frac{a_0}{a_n} z^{-n}) = a_n z^n (1 + g(z^{-1}))$, где $g(z^{-1})$ — многочлен от z^{-1} с нулевым свободным членом. Оценим снизу $|f(z)| = |a_n z^n||1 + g(z^{-1})| \geqslant |a_n z^n|\big|1 - |g(z^{-1})|\big|$. Здесь $|a_n z^n|$ сколь угодно большое при достаточно большом $|z|$, а по лемме 3.7.1 $|g(z^{-1})|$ сколь угодно малое, например, $|g(z^{-1})| < 1/2$ при достаточно малом $|z^{-1}|$ или при достаточно большом $|z|$. Тогда $|f(z)| > |a_n z^n|/2$ сколь угодно большое при достаточно большом $|z|$.

Получим нужную оценку. По лемме 3.7.1 для $\varepsilon = 1/2$ найдётся $\delta > 0$ такое, что при $|z^{-1}| < \delta$ выполнено неравенство $|g(z^{-1})| < 1/2$. Для произвольного $M > 0$ при $|z| > \sqrt[n]{\frac{2M}{|a_n|}}$ справедливо неравенство $|a_n z^n| > 2M$. Выберем $R = max\left(\sqrt[n]{\frac{2M}{|a_n|}}, \frac{1}{\delta}\right)$. Тогда если $|z| > R$, то $|z^{-1}| < \delta$ и $|z| > \sqrt[n]{\frac{2M}{|a_n|}}$, поэтому $|f(z)| = |a_n z^n||1 + g(z^{-1})| \geqslant |a_n z^n|\big|1 - |g(z^{-1})|\big| > 2M\left(1 - \frac{1}{2}\right) = M$. ■

Число $z_0 = x_0 + iy_0$ называется пределом последовательности $z_n = x_n + iy_n$, если для любого $\varepsilon > 0$ существует натуральное число n_0 такое, что $|z_n - z_0| < \varepsilon$ для всех n_0. Обозначение: $\lim\limits_{n\to\infty} z_n = z_0$.

Очевидно, что сходимость последовательности комплексных чисел $z_n = x_n + iy_n$ к $z_0 = x_0 + iy_0$ равносильна сходимости двух последовательностей действительных чисел: $\lim\limits_{n\to\infty} x_n = x_0$, $\lim\limits_{n\to\infty} y_n = y_0$.

Последовательность z_n называется ограниченной, если найдётся число $R > 0$ такое, что $|z_n| \leqslant R$ для всех n.

Лемма 3.7.5. (**Лемма Больцано-Вейерштрасса.**)

Из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность.

Доказательство леммы 3.7.5. Пусть $z_n = x_n + iy_n$ и $|z_n| \leqslant R$ для всех n. Тогда $|x_n| \leqslant |z_n| \leqslant R$ и $|y_n| \leqslant |z_n| \leqslant R$ для всех n. По лемме Больцано-Вейерштрасса из ограниченной вещественной последовательности x_n можно выделить сходящуюся подпоследовательность $x_{n_k} \to x_0$. Рассмотрим подпоследовательность y_{n_k}. Она ограничена и из неё можно выделить сходящуюся подпоследовательность $y_{n_{k_m}} \to y_0$. Тогда $x_{n_{k_m}} \to x_0$ и $z_{n_{k_m}} = x_{n_{k_m}} + iy_{n_{k_m}} \to z_0 = x_0 + iy_0$. ■

Лемма 3.7.6. (**Лемма о достижении инфимума модуля многочлена.**)

Точная нижняя грань модуля многочлена $f(z)$ достигается, т.е. существует число z_0 такое, что $|f(z_0)| \leqslant |f(z)|$ при всех $z \in C$.

Доказательство леммы 3.7.6. Рассмотрим множество значений модуля многочлена $\{|f(z)| \colon z \in C\}$. Поскольку $|f(z)| \geqslant 0$, то это множество ограничено снизу и, следовательно, имеет инфимум, который обозначим через m. Тогда для любого натурального числа n можно найти комплексное число z_n такое, что

$$m \leqslant |f(z_n)| \leqslant m + \frac{1}{n}. \tag{50}$$

Воспользуемся леммой 3.7.4 и для $M = m + 1$ найдём R такое, что при $|z| > R$ будет $|f(z)| > M$. Поэтому при $|z| > R$ выполнено неравенство $|f(z)| > m + \dfrac{1}{n}$ для всех n, т. к. $M = m + 1 \geqslant m + \dfrac{1}{n}$. Для z_n неравенство $|z_n| > R$ не возможно, т. к. тогда $|f(z_n)| > m + \dfrac{1}{n}$ вопреки (50). Следовательно, последовательность z_n ограничена. Согласно лемме 3.7.5 из z_n можно выделить сходящуюся подпоследовательность $z_{n_k} \to$ $\to z_0$. Тогда в силу непрерывности $|f(z)|$ (лемма 3.7.3)

$$\lim_{k\to\infty} |f(z_{n_k})| = |f(z_0)|. \tag{51}$$

Из неравенства (50) для подпоследовательности z_{n_k} получаем неравенство $m \leqslant |f(z_{n_k})| \leqslant m + \dfrac{1}{n_k}$, поэтому

$$\lim_{k\to\infty} |f(z_{n_k})| = m. \tag{52}$$

В силу единственности предела из (51) и (52) следует, что $|f(z_0)| = m$. ■

Лемма 3.7.7. (**Лемма Даламбера.**)

Если $f(z)$ — многочлен степени $n \geqslant 1$ и $f(z_0) \neq 0$, то найдётся число z_1 такое, что $|f(z_1)| < |f(z_0)|$.

Доказательство леммы 3.7.7. Для многочлена $f(z) = a_0 + a_1 z$ утверждение выполнено при $z_1 = -\dfrac{a_0}{a_1}$. При $n > 1$ разложим многочлен $f(z)$ по степеням $z - z_0$:

$$f(z) = c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n. \tag{53}$$

Очевидно, что $c_0 = f(z_0) \neq 0$. Пусть c_k — первый ненулевой коэффициент в (53) после c_0 (такой коэффициент имеется, так как $f(z)$ не константа). Преобразуем многочлен (53)

$$\begin{aligned} f(z) &= c_0 + c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \ldots + c_n(z - z_0)^n = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_{k+1}}{c_0}(z - z_0)^{k+1} + \ldots + \frac{c_n}{c_0}(z - z_0)^n\right) = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_k}{c_0}(z - z_0)^k\left(\frac{c_{k+1}}{c_k}(z - z_0) + \ldots + \frac{c_n}{c_k}(z - z_0)^{n-k}\right)\right) = \\ &= c_0\left(1 + \frac{c_k}{c_0}(z - z_0)^k + \frac{c_k}{c_0}(z - z_0)^k g(z - z_0)\right), \end{aligned} \tag{54}$$

где $g(z - z_0) = \dfrac{c_{k+1}}{c_k}(z - z_0) + \ldots + \dfrac{c_n}{c_k}(z - z_0)^{n-k}$ — многочлен от $z - z_0$ с нулевым свободным членом, предположим сначала, что $g(z - z_0)$ ненулевой. Введём обозначения $h = z - z_0$, $s = \dfrac{c_k}{c_0}$ и запишем равенство (54) в виде $f(z_0 + h) = c_0\left((1 + sh^k) + sh^k g(h)\right)$. Отсюда следует оценка $|f(z_0 + h)| = |c_0|\,|(1 + sh^k) + sh^k g(h)| \leqslant |c_0|(|1 + sh^k| + |sh^k||g(h)|)$. Если найдём такое $h \neq 0$, при котором $0 < |sh^k| < 1$, $|1 + sh^k| = 1 - |sh^k|$ и $0 < |g(h)| < \dfrac{1}{2}$, то $|f(z_0 + h)| < |c_0|(1 - |sh^k| + \dfrac{1}{2}|sh^k|) = |c_0|(1 - \dfrac{1}{2}|sh^k|) < |c_0| = |f(z_0)|$.

Будем искать h в виде $h = r(\cos\varphi + i\sin\varphi)$, $r = |h|$, $\varphi = \arg(h)$. Причём сначала выберем и фиксируем достаточно малое значение $r > 0$ такое, что при всех h, для которых $|h| = r$, выполнены неравенства $0 < |g(h)| < \dfrac{1}{2}$ и $0 < |sh^k| < 1$. Затем выберем $\varphi = \arg(h)$ так, чтобы sh^k являлось вещественным отрицательным числом. Для этого найдём такой $\arg(h)$, для которого $\arg(sh^k) = \pi$. Поскольку $\arg(sh^k) = \arg(s) + \arg(h^k) = \arg(s) + k\arg(h) = \arg(s) + k\varphi$ достаточно взять $\varphi = \varphi_1 = \dfrac{\pi - arg(s)}{k}$, чтобы для $h_1 = r(\cos\varphi_1 + i\sin\varphi_1)$ при выбранном r выполнялись условия $0 < |g(h_1)| < \dfrac{1}{2}$, $0 < |sh_1^k| < 1$, sh_1^k является вещественным отрицательным числом. Отсюда $-1 < sh_1^k < 0$, $-|sh_1^k| = sh_1^k$, $1 + sh_1^k > 0$, $|1 + sh_1^k| = 1 + sh_1^k = 1 - |sh_1^k|$. Тогда по ранее доказанному $|f(z_0 + h_1)| < |f(z_0)|$.

Доказательство теоремы 3.7.1. Пусть $f(z)$ — произвольный многочлен степени $n \geqslant 1$ над полем C от переменной z. Согласно лемме 3.7.6 множество всех значений $|f(z)|$ имеет инфимум m, который достигается в некоторой точке z_0, т.е. $|f(z_0)| = m$. Значение $m = 0$. Действительно, если $|f(z_0)| \neq 0$, то по лемме 3.7.7 найдётся число z_1 такое, что $|f(z_1)| < |f(z_0)| = \inf\{|f(z)|\}$, что невозможно. Таким образом, z_0 — корень многочлена $f(z)$. ■

3.8. Следствия из основной теоремы алгебры для многочленов над полем C

В этом разделе получим следующие следствия из основной теоремы алгебры для многочленов над C : существование и единственность разложения многочлена на линейные множители, каноническое разложение многочлена на множители, эквивалентность определений равенства двух многочленов по коэффициентам и по значениям, интерполяционный многочлен Лагранжа, формулы Виета.

Доказать следующие утверждения самостоятельно.

Теорема 3.8.1. (Существование и единственность разложения многочлена на линейные множители.)

Для любого многочлена $f(z) = \sum\limits_{k=0}^{n} a_k z^k \in C[z]$ *степени* $n \geqslant 1$ *существуют числа* $c_1, c_2, \ldots, c_n \in C$ *такие, что*

$$f(z) = a_n(z - c_1)(z - c_2)\ldots(z - c_n). \tag{55}$$

Это разложение единственно с точностью до порядка сомножителей.

Доказательство. I. Существование докажем ММИ по n.

1) При $n = 1$ имеем $f(z) = a_0 + a_1 z = a_1\Big(z + \dfrac{a_0}{a_1}\Big)$ и (55) выполнено для $z = -\dfrac{a_0}{a_1}$.

2) Пусть для произвольного многочлена $f(z) = \sum\limits_{k=0}^{m} a_k z^k$ степени m существуют числа $c_1, \ldots, c_m \in C$ такие, что $f(z) = a_m(z - c_1)\ldots(z - c_m)$.

3) Покажем, что утверждение верно для многочлена $f(z) = \sum\limits_{k=0}^{m+1} a_k z^k$ степени $m + 1$. Из основной теоремы 3.7.1 следует существование корня $z = c_{m+1}$ многочлена $f(z)$ и из следствия 1 теоремы Безу следует разложение $f(z) = h(z)(z - c_{m+1})$. При этом многочлен $h(z)$ имеет степень m и старший коэффициент a_{m+1}. Тогда по предположению индукции существуют числа $c_1, \ldots, c_m \in C$ такие, что $h(z) = a_{m+1}(z - c_1)\ldots(z - c_m)$. Следовательно, $f(z) = a_{m+1}(z - c_1)\ldots(z - c_m)(z - c_{m+1})$.

II. Докажем единственность разложения (55). Предположим, что есть второе разложение $f(z) = a_n(z - d_1)(z - d_2)\ldots(z - d_n)$. Тогда справедливо равенство $a_n(z - c_1)(z - c_2)\ldots(z - c_n) = a_n(z - d_1)(z - d_2)\ldots(z - d_n)$. Отсюда, каждый корень c_j совпадает с некоторым корнем d_k и каждый корень d_q совпадает с некоторым корнем c_r. Пусть произвольный корень $z = \alpha$ встречается в первом разложении s раз, а во втором разложении — t раз. Покажем, что $s = t$. $\sqcap$ $s > t$. Тогда $a_n(z - \alpha)^s \prod\limits_{c_j \neq \alpha} (z - c_j) = a_n(z - \alpha)^t \prod\limits_{d_q \neq \alpha} (z - d_q)$. После сокращения на $a_n(z - \alpha)^t$ получим равенство $(z - \alpha)^{s-t} \prod\limits_{c_j \neq \alpha} (z - c_j) = \prod\limits_{d_q \neq \alpha} (z - d_q)$, которое не выполнено при $z = \alpha$. $\sqcup$ ■

Теорема 3.8.2. (Каноническое разложение на множители многочлена над полем C.)

Для любого многочлена $f(z) = \sum\limits_{k=0}^{n} a_k z^k$ *степени* $n \geqslant 1$ *существуют число* $m \in N$, $1 \leqslant m \leqslant n$, *числа* $c_1, c_2, \ldots, c_m \in C$, $c_i \neq c_j$, *при* $i \neq j$, *и числа* $k_1, k_2, \ldots, k_m$, $k_1 + k_2 + \ldots + k_m = n$ *такие, что*

$$f(z) = a_n(z - c_1)^{k_1}(z - c_2)^{k_2} \ldots (z - c_m)^{k_m}. \tag{56}$$

Это разложение единственно с точностью до порядка сомножителей. Числа $c_1, c_2, \ldots, c_m$ *являются корнями многочлена* $f(z)$ *кратностей* $k_1, k_2, \ldots, k_m$.

Доказательство. Согласно теореме 3.8.1 справедливо разложение многочлена на множители $f(z) = a_n(z - d_1)(z - d_2) \ldots (z - d_n)$.

Если все корни d_j совпадают, то, обозначив их через c_1, получим разложение $f(z) = a_n(z - c_1)^n$.

Если среди корней d_j имеется $m\,(1 < m \leqslant n)$ различных d_{j_1}, $d_{j_2}, \ldots, d_{j_m}$, которые встречаются $k_1, k_2, \ldots, k_m$ раз соответственно, то введём обозначения $c_1 = d_{j_1}, c_2 = d_{j_2}, \ldots, c_m = d_{j_m}$ и получим разложение $f(z) = a_n(z - c_1)^{k_1}(z - c_2)^{k_2} \ldots (z - c_m)^{k_m}$. ■

Следствие. Любой многочлен $f(z) \in C[z]$ степени $n \geqslant 1$ имеет ровно n корней, если каждый корень считать столько раз, какова его кратность. ■

Замечание. Это утверждение применимо к многочлену нулевой степени, т.к. такой многочлен не имеет корней. Это утверждение неприменимо к нулевому многочлену, т.к. такой многочлен имеет корнем любое число $z \in C$. ■

Рассмотрим вопрос о эквивалентности определений равенства двух многочленов по коэффициентам и по значениям.

Теорема 3.8.3. (Равенство двух многочленов.)

Если два многочлена $f(z), g(z) \in C[z]$, *степени которых не превосходят* $n \in N$, *имеют равные значения при более чем* n *различных значениях переменной, то они равны.*

Доказательство. Многочлен $h(z) = f(z) - g(z)$ имеет степень не выше n и более n различных корней. Поэтому $h(z)$ — нулевой многочлен и тогда $f(z) = g(z)$.

Следствия. 1. *Два многочлена* $f(z)$, $g(z) \in C[z]$ *равны тогда и только тогда, когда совпадают их значения при всех значениях переменной* z.

2. *Многочлен* $f(z) \in C[z]$ *степени не выше* n *однозначно определяется своими значениями в* $n+1$ *различных точках.*

Доказательство. 1. а) Если многочлены $f(z) = \sum\limits_{k=0}^{\infty} f_k z^k$ и $g(z) = \sum\limits_{k=0}^{\infty} g_k z^k$ равны, то равны все их коэффициенты $f_k = g_k$, $k \in Z_+$, и тогда совпадают их значения при всех значениях переменной. Причём это верно для многочленов над любым полем.

б) Если значения многочленов $f(z)$, $g(z) \in C[z]$ совпадают при всех значениях переменной z, то по теореме 3.8.3 они равны.

2. Если значения многочленов $f(z)$, $g(z) \in C[z]$, степени не выше n, совпадают для $n+1$ различных точек, то по теореме 3.8.3 они равны. ■

Теорема 3.8.4. (Интерполяционный многочлен.)

Для любых чисел $z_1, z_2, \ldots, z_{n+1} \in C$, $z_i \neq z_j$ *при* $i \neq j$ *и любых чисел* $u_1, u_2, \ldots, u_{n+1} \in C$ *существует единственный многочлен* $f(z) \in C[z]$ *степени не выше* n, *для которого выполнены условия интерполяции*

$$f(z_k) = u_k,\ k = \overline{1, n+1}. \tag{57}$$

Этот многочлен можно записать в форме Лагранжа

$$f(z) = \sum_{k=1}^{n+1} u_k \frac{(z-z_1)\ldots(z-z_{k-1})(z-z_{k+1})\ldots(z-z_{n+1})}{(z_k-z_1)\ldots(z_k-z_{k-1})(z_k-z_{k+1})\ldots(z_k-z_{n+1})} =$$
$$= \sum_{k=1}^{n+1} u_k \prod_{j\neq k} \frac{(z-z_j)}{(z_k-z_j)} \tag{58}$$

или в форме Ньютона

$$f(z) = v_1 + v_2(z-z_1) + v_3(z-z_1)(z-z_2) + \ldots +$$
$$+ v_{n+1}(z-z_1)(z-z_2)\ldots(z-z_n), \tag{59}$$

где v_j, $j = \overline{1, n+1}$ *последовательно вычисляются с использованием условий интерполяции (57).*

Доказательство. Формула (58) определяет многочлен степени не выше n. При $z = z_k$ все слагаемые, кроме k-го, обращаются в 0, а

k-е слагаемое $u_k \dfrac{(z_k - z_1)\dots(z_k - z_{k-1})(z_k - z_{k+1})\dots(z_k - z_{n+1})}{(z_k - z_1)\dots(z_k - z_{k-1})(z_k - z_{k+1})\dots(z_k - z_{n+1})} = u_k.$
Поэтому условия интерполяции $f(z_k) = u_k$, $k = \overline{1, n+1}$ выполнены, и, согласно следствию 2 теоремы 3.8.3, такой многочлен единственен.

Теорема 3.8.5. (Формулы Виета.)

Если $c_1, c_2, \dots, c_n \in C$ *— корни многочлена* $f(z) = \sum\limits_{k=0}^{n} a_k z^k \in C[z]$ *степени* $n \geqslant 1$, *то*

$$\begin{aligned}
&a_{n-1}/a_n = -(c_1 + c_2 + \dots + c_n),\\
&a_{n-2}/a_n = +(c_1c_2 + c_1c_3 + \dots + c_1c_n + c_2c_3 + \dots + c_{n-1}c_n),\\
&a_{n-3}/a_n = (-1)^3 \sum_{i_1<i_2<i_3} c_{i_1}c_{i_2}c_{i_3},\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&a_{n-k}/a_n = (-1)^k \sum_{i_1<i_2<\dots<i_k} c_{i_1}c_{i_2}\dots c_{i_k},\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&a_0/a_n = (-1)^n c_1c_2\dots c_n.
\end{aligned} \tag{60}$$

Доказательство. Согласно теореме 3.8.1 имеем

$$\sum_{k=0}^{n} a_k z^k = a_n(z - c_1)(z - c_2)\dots(z - c_n) \Leftrightarrow$$

$$\Leftrightarrow \sum_{k=0}^{n} \frac{a_k}{a_n} z^k = (z - c_1)(z - c_2)\dots(z - c_n).$$

Приравняем коэффициенты при степени z^{n-k} в левой и правой частях равенства. Для этого в правой части равенства из суммы 2^n произведений нужно выбрать те C_n^k, которые содержат $n-k$ множителей z и k множителей, выбранных из $(-c_1)$, $(-c_2)$, $\dots$, $(-c_n)$. Получим

$$\frac{a_{n-k}}{a_n} = \sum_{i_1<i_2<\dots<i_k} (-c_{i_1})(-c_{i_2})\dots(-c_{i_k}) \Leftrightarrow$$

$$\Leftrightarrow \frac{a_{n-k}}{a_n} = (-1)^k \sum_{i_1<i_2<\dots<i_k} c_{i_1}c_{i_2}\dots c_{i_k}. \blacksquare$$

3.9. Многочлены над полем R

Пусть $f(x) \in R[x]$ — многочлен с вещественными коэффициентами. Так как $R \subset C$, то $f(x)$ можно рассматривать как многочлен над полем

C и тогда для этого многочлена справедливы все утверждения из 3.8. Уточним некоторые из этих утверждений для многочленов над R.

Теорема 3.9.1. (О сопряжённых корнях многочлена над R.)

Если $c = \alpha + \beta i\,(\beta \neq 0)$ — комплексный корень ненулевого многочлена из $R[x]$, то $\overline{c} = \alpha - \beta i$ также является корнем этого многочлена, причём кратности корней c и $\overline{c}$ одинаковы.

Доказательство. I. Сначала докажем следующее утверждение: *комплексное число c является корнем многочлена с вещественными коэффициентами тогда и только тогда, когда сопряжённое число $\overline{c}$ является корнем этого многочлена.*

Если комплексное число c является корнем многочлена $f(x) = \sum\limits_{k=0}^{n} a_k x^k \in R[x]$, то $f(c) = 0 \Leftrightarrow a_0 + a_1 c + \ldots + a_n c^n = 0$. Возьмём сопряжение от обеих частей равенства и преобразуем:

$$f(c) = 0 \Leftrightarrow \overline{f(c)} = \overline{0} \Leftrightarrow \overline{a_0 + a_1 c + \ldots + a_n c^n} = \overline{0} \Leftrightarrow$$
$$\Leftrightarrow \overline{a}_0 + \overline{a_1 c} + \ldots + \overline{a_n c^n} = \overline{0} \Leftrightarrow \overline{a}_0 + \overline{a}_1 \overline{c} + \ldots + \overline{a}_n \overline{c}^n = \overline{0} \Leftrightarrow$$
$$\Leftrightarrow a_0 + a_1 \overline{c} + \ldots + a_n \overline{c}^n = 0 \Leftrightarrow f(\overline{c}) = 0.$$

При условии $\beta \neq 0$ корни $c = \alpha + \beta i$ и $\overline{c} = \alpha - \beta i$ различны.

Отсюда следует, что *для любого многочлена из $R[x]$ справедливо только одно из условий c и $\overline{c}$ не являются корнями этого многочлена или c и $\overline{c}$ являются корнями этого многочлена.*

II. Теперь докажем, что *если c и $\overline{c}$ — различные корни многочлена $f(x) \in R[x]$, то справедливо разложение $f(x) = (x - c)(x - \overline{c}) f_1(x)$, где $f_1(x) \in R[x]$.*

Поскольку $x = c$ и $x = \overline{c}$ различные корни многочлена $f(x)$, то в кольце многочленов $C[x]$, согласно следствию из теоремы Безу, $f(x)$ делится на каждый из многочленов $x - c$ и $x - \overline{c}$. При $c \neq \overline{c}$ общими делителями $x - c$ и $x - \overline{c}$ являются только ненулевые константы, т. е. многочлены взаимно просты. Тогда по теореме 3.5.2 $f(x)$ делится на произведение $\varphi(x) = (x - c)(x - \overline{c})$, т. е. имеем $f(x) = \varphi(x) f_1(x)$, где $f_1(x) \in C[x]$. Покажем, что $f_1(x) \in R[x]$. После раскрытия скобок $(x - c)(x - \overline{c}) = x^2 - (c + \overline{c})x + c\overline{c} = x^2 - 2\alpha x + (\alpha^2 + \beta^2)$ видим, что $\varphi(x) \in R[x]$. Из теоремы о делении многочленов в кольце

$R[x]$ следует, что для $f(x), \varphi(x) \in R[x]$ существуют единственная пара многочленов $u(x), v(x) \in R[x]$ такая, что $f(x) = \varphi(x)u(x) + v(x)$ и $\deg(v(x)) < \deg(\varphi(x))$. Равенство $f(x) = \varphi(x)u(x)+v(x)$ можно рассматривать как результат деления $f(x)$ на $\varphi(x)$ в кольце $C[x]$. Но в кольце $C[x]$ справедливо равенство $f(x) = \varphi(x)f_1(x)$. Поэтому в силу единственности результата деления имеем $f_1(x) = u(x)$ и $v(x) = 0$. Следовательно, $f_1(x) \in R[x]$ и выполнено равенство $f(x) = (x - c)(x - \overline{c})f_1(x)$.

III. Теперь докажем, что *кратности корней c и $\overline{c}$ ненулевого многочлена с вещественными коэффициентами совпадают.*

Из условия теоремы и утверждения части **I** следует, что многочлен $f(x) \in R[x]$ имеет корни c и $\overline{c}$ $(c \neq \overline{c})$. Согласно части **II**, верно разложение $f(x) = (x - c)(x - \overline{c})f_1(x)$, где $f_1(x) \in R[x]$.

Если для $f_1(x)$ выполнено условие $f_1(c) \neq 0$, $f_1(\overline{c}) \neq 0$, то выделение множителей $(x - c)(x - \overline{c})$ завершено и $f(x)$ имеет корни c, $\overline{c}$ кратности 1.

Если для $f_1(x)$ выполнено условие $f_1(c) = 0$, $f_1(\overline{c}) = 0$, то продолжим выделение множителей $(x - c)(x - \overline{c})$ и, согласно части **II**, получим $f_1(x) = (x - c)(x - \overline{c})f_2(x)$, где $f_2(x) \in R[x]$. При этом будет верно разложение $f(x) = (x - c)^2(x - \overline{c})^2 f_2(x)$, где $f_2(x) \in R[x]$.

Если для $f_2(x)$ выполнено условие $f_2(c) \neq 0$, $f_2(\overline{c}) \neq 0$, то выделение множителей $(x - c)(x - \overline{c})$ завершено и $f(x)$ имеет корни c, $\overline{c}$ кратности 2. Если для $f_2(x)$ выполнено условие $f_2(c) = 0$, $f_2(\overline{c}) = 0$, то продолжим выделение множителей $(x - c)(x - \overline{c})$.

Повторяя процесс выделения множителей $(x - c)(x - \overline{c})$ k раз $\left(k \leqslant \dfrac{1}{2}\deg(f(x))\right)$, получим разложение $f(x) = (x - c)^k(x - \overline{c})^k f_k(x)$, для которого $f_k(x) \in R[x]$, $f_k(c) \neq 0$, $f_k(\overline{c}) \neq 0$. Тогда $f(x)$ имеет корни c, $\overline{c}$ кратности k.

Следствие. Любой многочлен $f(x) \in R[x]$ нечётной степени имеет хотя бы один вещественный корень.

Теорема 3.9.2. (Каноническое разложение на множители многочлена над полем $\mathbf{R}$.)

Для любого многочлена $f(x) = \sum\limits_{k=0}^{n} a_k x^k \in R[x]$ *степени* $n \geqslant 1$

справедливо разложение

$$f(x) = a_n \prod_{j=1}^{m} (x - c_j)^{k_j} \prod_{s=1}^{r} (x^2 + p_s x + q_s)^{t_s}. \tag{61}$$

где числа $m, r \in Z_+$, $m + r > 0$, *при* $m \neq 0$ *числа* $c_1, \ldots, c_m$ *— это все различные вещественные корни данного многочлена с кратностями* $k_1, \ldots, k_m \in N$, *при* $r \neq 0$ *квадратные трёхчлены* $x^2 + p_1 x + q_1, \ldots, x^2 + p_r x + q_r$ *различны, имеют вещественные коэффициенты и отрицательные дискриминанты, а показатели степеней этих трёхчленов* $t_1, \ldots, t_m$ *являются натуральным числами, кроме того выполнено равенство* $\sum\limits_{j=1}^{m} k_j + 2 \sum\limits_{s=1}^{r} t_s = n$.

Разложение (61) единственно с точностью до порядка сомножителей.

Часть II
Линейные пространства. Линейные операторы

4. Линейное пространство над произвольным полем

4.1. Линейное пространство над произвольным полем, его простейшие свойства. Подпространство. Критерий подпространства

Определение 1.

I. Пусть задано множество $E \neq \emptyset$, *элементы которого называем векторами, и поле* P, *элементы которого называем числами.*

II. Пусть определены две операции

а) сложение векторов: $\forall f, g \in E \ \exists!\, h \in E$ *обозначаемый* $h = f + g$ *и называемый суммой векторов* f *и* g,

б) умножение вектора на число: $\forall \alpha \in P, f \in E \ \exists!\, w \in E$, *обозначаемый* αf *и называемый произведением вектора* f *на число* α.

III. Пусть выполняются следующие аксиомы:

АЛП1 $\forall f, g, h \in E \quad (f + g) + h = f + (g + h)$ *— ассоциативность сложения;*

АЛП2 $\exists\, \theta \in E \ \forall f \in E \quad f + \theta = f$, θ *называется нулевым вектором;*

АЛП3 $\forall f \in E \, \exists\, g \in E \quad f + g = \theta$, g *называется противоположным для вектора* f *и обозначается* $(-f)$;

АЛП4 $\forall f, g \in E \quad f + g = g + f$ *— коммутативность сложения;*

АЛП5 $\forall \alpha, \beta \in P, f \in E \quad \alpha(\beta f) = (\alpha\beta) f$;

АЛП6 $\forall f \in E \quad 1f = f$;

АЛП7 $\forall \alpha, \beta \in P, f \in E \quad (\alpha + \beta) f = \alpha f + \beta f$ *— дистрибутивность умножения вектора на сумму чисел;*

АЛП8 $\forall \alpha \in P, f, g \in E \quad \alpha(f+g) = \alpha f + \alpha g$ *— дистрибутивность*

умножения суммы векторов на число.

Тогда E называется линейным пространством над полем P, операции а) и б) называются линейными операциями.

Если $P = R$ ($P = C$), то E называется вещественным (комплексным) линейным пространством. ■

Замечания. *В силу аксиом 2 – 4 выполнены следующие свойства.*

Для любого вектора $f \in E$ $\theta + f = f$, $f + \theta = f$.

Если вектор $(-f) \in E$ является противоположным для $f \in E$, то $(-f) + f = \theta$, т.е. $f = -(-f)$, вектор f является противоположным для вектора $(-f)$.

Примеры. Множество всех геометрических векторов V_3 над полем R. Используются операции сложения векторов и умножения векторов на вещественные числа. Векторы образуют вещественное линейное пространство.

Множество вещественных чисел R (это множество векторов) над полем R. Используются операции сложения векторов и умножения векторов на числа из поля R. Векторы образуют вещественное линейное пространство.

Множество комплексных чисел C (это множество векторов) над полем C. Используются операции сложения векторов и умножения векторов на числа из поля C. Векторы образуют комплексное линейное пространство.

Множество комплексных чисел C (это множество векторов) над полем R. Используются операции сложения векторов и умножения векторов на числа из поля R. Векторы образуют вещественное линейное пространство. ■

Теорема 4.1.1 (Простейшие свойства пространства).

Для произвольного линейного пространства E над произвольным полем P выполнены следующие свойства.

1. *Нулевой вектор единственен.*

2. *Для любого вектора противоположный к нему вектор единственен.*

3. *Для любых векторов* $f, g, h \in E$ *из* $f+h=g+h$ *или* $h+f=h+g$ *следует* $f=g$.

4. *Для любых векторов* $f, g \in E$ *существует, и притом единственное, решение* $x=g+(-f)$ *уравнения* $x+f=g$; *это решение называется разностью* g *и* f *и обозначается* $g-f$. *Итак,* $g-f=g+(-f)$ *и эта формула определяет операцию вычитания.*

5. $\forall \alpha \in P, f, g \in E \quad \alpha(f-g)=\alpha f-\alpha g$.

6. $\forall \alpha, \beta \in P, f \in E \quad (\alpha-\beta)f=\alpha f-\beta f$.

7. $\forall\, \alpha \in P \quad \alpha\theta=\theta$.

8. $\forall\, f \in E \quad 0f=\theta$.

9. *Из равенства* $\alpha f=\theta$, $\alpha \in P$, $f \in E$ *следует* $\alpha=0$ *или* $f=\theta$.

10. $\forall f \in E \quad (-f)=(-1)f$ (*противоположный вектор для* f *равен* $(-1)f$).

Доказательство.

1. Пусть θ_1 и θ_2 нулевые векторы. Тогда $\theta_1+\theta_2=\theta_1$, т.к. θ_2 нулевой вектор, и $\theta_1+\theta_2=\theta_2$, т.к. θ_1 нулевой вектор. Отсюда $\theta_1=\theta_2$.

2. Пусть g_1 и g_2 — противоположные векторы для f, т.е. $g_1+f=\theta$, $f+g_2=\theta$. Из ассоциативности сложения следует, что $(g_1+f)+g_2=g_1+(f+g_2)$. Отсюда получим равенства $\theta+g_2=g_1+\theta$ и $g_2=g_1$.

3. Прибавим к обеим частям равенства $f+h=g+h$ вектор $(-h)$. Получим $(f+h)+(-h)=(g+h)+(-h)$. Отсюда следуют равенства $f+(h+(-h))=g+(h+(-h))$, $f+\theta=g+\theta$ и $f=g$.

4. Проверим, что $x=g+(-f)$ решение уравнения $x+f=g$. Действительно, $x+f=g+(-f)+f=g+\theta=g$.

Пусть $x=c$ решение уравнения $x+f=g$, т.е. $c+f=g$. Прибавим к обеим частям равенства $c+f=g$ элемент $(-f)$. Получим $c+f+(-f)=$ $=g+(-f)$. Отсюда следуют равенства $c+\theta=g+(-f)$ и $c=g+(-f)$, т.е. решение единственно.

5. Для разности $f-g$ верно равенство $(f-g)+g=f$. Умножим обе части равенства на α, получим $\alpha((f-g)+g)=\alpha f$ и $\alpha(f-g)+\alpha g=\alpha f$. Отсюда по определению разности имеем $\alpha(f-g)=\alpha f-\alpha g$.

6. Для разности чисел $\alpha - \beta$ верно равенство $(\alpha - \beta) + \beta = \alpha$. Умножим вектор f на это число, получим $((\alpha-\beta)+\beta)f = \alpha f$ и $(\alpha-\beta)f+\beta f = \alpha f$. Отсюда по определению разности имеем $(\alpha - \beta)f = \alpha f - \beta f$.

7. Из свойства 5 при $g = f$ получаем $\alpha(f - f) = \alpha f - \alpha f$. Отсюда в силу $f - f = \theta$ и $\alpha f - \alpha f = \theta$ следует равенство $\alpha\theta = \theta$.

8. Из свойства 6 при $\beta = \alpha$ получаем $(\alpha - \alpha)f = \alpha f - \alpha f$. Отсюда в силу $\alpha - \alpha = 0$ и $\alpha f - \alpha f = \theta$ следует равенство $0f = \theta$.

9. Если $\alpha = 0$, то утверждение верно. При $\alpha \neq 0$ существует обратное число α^{-1}. Умножим равенство $\alpha f = \theta$ на α^{-1}, получим $\alpha^{-1}\alpha f = \alpha^{-1}\theta$. Отсюда следует равенство $1f = \theta$, т.е. $f = \theta$.

10. Сложим векторы $f = 1f$ и $(-1)f$. Получим $f + (-1)f = 1f + (-1)f = (1 + (-1))f = 0f = \theta$, т.е. $f + (-1)f = \theta$. Следовательно, вектор $(-1)f$ является противоположным к f, т.е. $(-f) = (-1)f$. ■

Определение 2. *Пусть E — линейное пространство над полем P и L — непустое подмножество в E.*

L называется линейным подпространством линейного пространства E или подпространством E, если L само является линейным пространством над полем P для тех же операций как и E. ■

Теорема 4.1.2 (Критерий линейного подпространства).

Подмножество $L \neq \emptyset$ линейного пространства E над полем P является подпространством тогда и только тогда, когда L замкнуто относительно линейных операций, т.е.

а) $\forall f, g \in L \quad f + g \in L$;

б) $\forall f \in L, \alpha \in P \quad \alpha f \in L$.

Доказательство. I. Пусть L - подпространство E. Тогда в L определены такие же как и в E операции сложения векторов и умножения векторов на числа из P. Тогда выполнены свойства а) и б).

II. Пусть в L выполнены свойства а) и б), следовательно, определены операции сложения векторов и умножения векторов на числа из P. Тогда в L выполнены аксиомы 1, 4, 5, 6, 7, 8, так как равенства в аксиомах выполнены в E, а $L \subset E$, причём L замкнуто относительно линейных операций.

Поскольку при $f \in L$ в силу свойства б) заключаем, что нулевой вектор $\theta = 0f \in L$, следовательно, аксиома 2 выполнена в L. Аналогично, для $f \in L$ в силу свойства б) заключаем, что противоположный вектор $(-1)f \in L$, и тогда аксиома 3 выполнена в L. ■

Задачи.

Задача 4.1.1. *Проверьте, является ли линейным пространством над полем P данное множество E с указанными операциями.*

а). Поле P, $E = P^{m\times n}$ — матрицы с элементами из поля P, операции сложения матриц и умножения матриц на числа из поля.

б). Поле P, $E = \{f \mid f\colon M \to P\}$ — функции, определённые на множестве M со значениями в поле P, операции сложения функций и умножения функций на числа из P. Для функций $f, g \in E$ сумма $f + g$ определяется $\forall x \in M$ равенством $(f + g)(x) = f(x) + g(x)$. Для функции $f \in E$ и числа $\alpha \in P$ произведение αf (функции f на число α) определяется равенством $(\alpha f)(x) = \alpha f(x)$, $x \in M$.

в). Поле P, $E = P[x]$ — многочлены от одной переменной x с коэффициентами из поля P, операции сложения многочленов и умножения многочленов на числа из поля P.

г). Поле $P = R$, $E = Z^{m\times n}$ — матрицы размера $m\times n$ из целых чисел, операции сложения матриц и умножения матриц на числа из R.

д). Поле $P = R$, E — рациональные дроби вида $\dfrac{f(x)}{g(x)}$, $f(x), g(x) \in R[x]$, $g(x) \not\equiv 0$, операции умножения дробей на числа из R $\alpha\dfrac{f(x)}{g(x)} = \dfrac{\alpha f(x)}{g(x)}$ и сложения дробей $\dfrac{f(x)}{g(x)} + \dfrac{h(x)}{k(x)} = \dfrac{f(x)k(x) + g(x)h(x)}{g(x)k(x)}$.

Задача 4.1.2. *Проверьте, является ли подмножество L линейным подпространством линейного пространства E.*

а). $E = R[x]$ — линейное пространство многочленов от одной переменной x с вещественными коэффициентами и с операциями сложения многочленов и умножения многочленов на вещественные числа, $L = R_2[x]$ — множество многочленов $f \in E:\ \deg(f) \leqslant 2$.

б). $E = R[x]$, L — *множество многочленов* $f \in E:\ f(1) = 0,\ f(2) = 1$.

в). $E = R^{2\times 3}$, L — *множество матриц* $A = (a_{ij}) \in E : a_{13} = a_{23} = 0$.

г). $E = R^{2\times 2}$, L — *множество матриц* $A = (a_{ij}) \in E : a_{11} + a_{22} = 1$.

д). $E = R^{2\times 2}$, L — *множество матриц* $A = (a_{ij}) \in E : a_{11} + 2a_{22} = 0,\ 2a_{12} - 3a_{21} = 0$.

е). $E = R[x]$, $L = R_2[x]$ — *множество многочленов* $f \in E : \deg(f) = 2$.

ж). $E = R[x]$, $L = \{f \in E : f(x) > 0$ *при* $x > 0\}$.

з). $E = R^{3\times 3}$, L — *подмножество симметрических матриц.*

и). $E = R^{3\times 3}$, L — *подмножество кососимметрических матриц.*

к). $E = R^{n\times n}$, L — *подмножество кососимметрических матриц, у которых последняя строка нулевая.*

Решение задач.

4.1.1 (а) I. Заданы поле P и множество векторов $E = P^{m\times n}$ — матрицы размера $m \times n$ вида $\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ с элементами a_{ij} из P.

II. 1) Сумма матриц $A = (a_{ij}) \in E$ и $B = (b_{ij}) \in E$ по определению равна $m \times n$-матрице $A + B = (a_{ij} + b_{ij})$. При этом $\forall\, i, j\ \ a_{ij} + b_{ij} \in P$. Отображение $\langle A, B\rangle \mapsto A + B$ определяет операцию сложения матриц.

2) Произведение матрицы $A = (a_{ij}) \in E$ на число $\beta \in P$ по определению равно $m \times n$-матрице $\beta A = (\beta a_{ij})$. При этом $\forall\, i, j\ \ \beta a_{ij} \in P$. Отображение $\langle \beta, A\rangle \mapsto \beta A$ определяет операцию умножения A на β.

III. Проверим аксиомы линейного пространства.

АЛП1 $\forall\, A, B, C \in E \quad (A + B) + C = A + (B + C)$.

Пусть $A = (a_{ij}) \in E$, $B = (b_{ij}) \in E$, $C = (c_{ij}) \in E$, тогда $(A + B) + C = \Big((a_{ij} + b_{ij}) + c_{ij}\Big)$, $A + (B + C) = \Big(a_{ij} + (b_{ij} + c_{ij})\Big)$ и $(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij})$ при всех i, j.

Следовательно, $(A+B)+C = A+(B+C)$ верно.

АЛП2 $\exists\theta \in E\ \forall A \in E \quad A+\theta = A$, верно, если θ — нулевая матрица.

АЛП3 $\forall A \in E\, \exists B \in E \quad A+B=\theta$ верно, если матрица B состоит из элементов противоположных элементам матрицы A.

АЛП4 $\forall A, B \in E \quad A+B = B+A$.

Пусть $A = (a_{ij}) \in E$, $B = (b_{ij}) \in E$, тогда $A+B = (a_{ij}+b_{ij})$, $B+A = (b_{ij}+a_{ij})$ и $a_{ij}+b_{ij} = b_{ij}+a_{ij}$ при всех i, j.

Следовательно, $A+B = B+A$ верно.

АЛП5 $\forall\alpha,\, \beta \in P,\, A \in E \quad \alpha(\beta A) = (\alpha\beta)A$;

Пусть $A = (a_{ij}) \in E$, тогда $\beta A = (\beta a_{ij}) \in E$, $\alpha(\beta A) = \Big(\alpha(\beta a_{ij})\Big) = \Big((\alpha\beta)a_{ij}\Big)$, $(\alpha\beta)A = \Big((\alpha\beta)a_{ij}\Big)$. Следовательно, $\alpha(\beta A) = (\alpha\beta)A$ верно.

АЛП6 $\forall A \in E \quad 1A = A$;

Пусть $A = (a_{ij}) \in E$, тогда $1A = (1a_{ij}) = (a_{ij}) = A$.

АЛП7 $\forall\alpha,\, \beta \in P,\, A \in E \quad (\alpha+\beta)A = \alpha A + \beta A$.

Пусть $A = (a_{ij}) \in E$, тогда $(\alpha+\beta)A = \Big((\alpha+\beta)a_{ij}\Big)$, $\alpha A = \Big(\alpha a_{ij}\Big)$, $\beta A = \Big(\beta a_{ij}\Big)$, $\alpha A + \beta A = \Big(\alpha a_{ij} + \beta a_{ij}\Big) = \Big((\alpha+\beta)a_{ij}\Big)$.

Следовательно, $(\alpha+\beta)A = \alpha A + \beta A$ верно.

АЛП8 $\forall\alpha \in P,\, A, B \in E \quad \alpha(A+B) = \alpha A + \alpha B$.

Пусть $A = (a_{ij}) \in E$, $B = (b_{ij}) \in E$, тогда $A+B = (a_{ij}+b_{ij})$, $\alpha(A+B) = \Big(\alpha(a_{ij}+b_{ij})\Big) = \Big(\alpha a_{ij} + \alpha b_{ij}\Big)$, $\alpha A = (\alpha a_{ij})$, $\alpha B = (\alpha b_{ij})$, $\alpha A + \alpha B = \Big(\alpha a_{ij} + \alpha b_{ij}\Big)$. Следовательно, $\alpha(A+B) = \alpha A + \alpha B$ верно.

Аксиомы линейного пространства выполнены. Тогда из I – III следует, что E — линейное пространство над полем P. ■

4.1.2 (a) Проверим с помощью критерия подпространства, является ли L подпространством E. Подмножество запишем в виде

$L = \{f \in R[x] \colon f(x) = a_0 + a_1x + a_2x^2,\, a_i \in R,\, i = \overline{0,2}\}$.

1) $f,\, g \in L \Rightarrow f+g \in L$.

$f,\, g \in L \Rightarrow \begin{cases} f(x) = a_0 + a_1x + a_2x^2,\ a_i \in R,\ i = \overline{0,2}, \\ g(x) = b_0 + b_1x + b_2x^2,\ b_i \in R,\ i = \overline{0,2}, \end{cases} \Rightarrow f(x) + g(x) =$
$= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2,\ a_i + b_i \in R,\ i = \overline{0,2} \Rightarrow f + g \in L.$

2) $\beta \in R,\ f \in L \Rightarrow \beta f \in L.$

$\beta \in R,\ f \in L \Rightarrow \beta \in R,\ f(x) = a_0 + a_1x + a_2x^2,\ a_i \in R,\ i = \overline{0,2} \Rightarrow$
$\Rightarrow \beta f(x) = (\beta a_0) + (\beta a_1)x + (\beta a_2)x^2,\ \beta a_i \in R,\ i = \overline{0,2} \Rightarrow \beta f \in L.$

Из 1) и 2) следует, что $R_2[x]$ является подпространством $R[x]$. ■

4.1.2 (б).

Подмножество запишем в виде $L = \{f \in R[x] \colon f(1) = 0,\ f(2) = 1\}$.

1) $f,\, g \in L \Rightarrow f + g \in L.$

Пусть $f,\, g \in L$. Тогда $f(1) = 0,\ f(2) = 1,\ g(1) = 0,\ g(2) = 1 \Rightarrow$ $\Rightarrow f(1) + g(1) = 0,\ f(2) + g(2) = 2$. Поскольку $f(2) + g(2) \neq 1$, то $f + g \notin$ $\notin L$. Так как критерий не выполнен, то L не является подпространством пространства $R[x]$. ■

4.2. Линейная комбинация векторов. Линейная зависимость и линейная независимость векторов

В этом разделе будем рассматривать зависимые и независимые системы векторов линейного пространства E над полем P.

Определение 1. *Пусть заданы векторы из E и числа из P :*

$$e_1,\, e_2,\, \ldots,\, e_k \in E, \tag{62}$$

$$\alpha_1,\, \alpha_2,\, \ldots,\, \alpha_k \in P. \tag{63}$$

$$\textit{Вектор} \quad \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = \sum_{i=1}^{k} \alpha_i e_i \tag{64}$$

называется линейной комбинацией векторов (62) *с коэффициентами* (63).

Линейная комбинация (64) *называется тривиальной, если все её коэффициенты равны* 0, *и нетривиальной, если в ней существует хотя бы один ненулевой коэффициент.*

Если вектор x есть линейная комбинация (64), *то говорят, что вектор x линейно выражается через векторы* (62), *а равенство* $x = \sum\limits_{i=1}^{k} \alpha_i e_i$ *называют разложением вектора x по векторам* (62). ■

Замечания.

1. Тривиальная линейная комбинация равна нулевому вектору θ.

2. Линейная комбинация одного вектора e_1 с коэффициентом α_1 — это произведение $\alpha_1 e_1$.

Определение 2. *Векторы $e_1, e_2, \ldots, e_k$ называются линейно зависимыми, если найдётся нетривиальная линейная комбинация этих векторов, равная θ, т. е. существуют числа $\alpha_1, \alpha_2, \ldots, \alpha_k \in P$, не все равные 0, и такие, что $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$.*

Векторы $e_1, e_2, \ldots, e_k$ называется линейно независимыми, если только тривиальная линейная комбинация этих векторов равна θ, т. е. если из условия $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$ следует $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

Также говорят, что система векторов линейно зависима (независима).

Теорема 4.2.1. 1° *При изменении порядка векторов их линейная зависимость (независимость) сохраняется.*

2° *Если часть векторов линейно зависима, то и все векторы линейно зависимы.*

Доказательство. 1° Если векторы $e_1, e_2, \ldots, e_k$ линейно зависимы, то существует нетривиальная линейная комбинация, равная θ : $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = \theta$. Так как при изменении порядка слагаемых сумма не меняется, то любая перестановка этих векторов образует линейно зависимую систему векторов. Следовательно, при изменении порядка векторов их линейная зависимость сохраняется.

Пусть данные векторы линейно независимы. Докажем «от противного», что при изменении порядка векторов их линейная независимость сохраняется.

$\sqcap$ Переставленный набор этих векторов образует линейно зависимую систему. Тогда, по доказанному выше, любая перестановка векторов, в том числе и в исходном порядке, образует линейно зависимую систему. $\sqcup$ Следовательно, при изменении порядка векторов их линейная независимость сохраняется.

2° Пусть среди k векторов m векторов линейно зависимы. Переставив векторы (при необходимости), считаем, что из k векторов $e_1, \ldots, e_m, e_{m+1}, \ldots, e_k$ линейно зависимы первые m векторов. Тогда $\sum\limits_{i=1}^{m} \alpha_i e_i = \theta$ для некоторых чисел α_i, $i = \overline{1, m}$, среди которых есть $\alpha_j \neq 0$. Добавив вектор $\sum\limits_{i=m+1}^{k} 0 e_i = \theta$, получим $\sum\limits_{i=1}^{m} \alpha_i e_i + \sum\limits_{i=m+1}^{k} 0 e_i = \theta$, где $\alpha_j \neq 0$. Следовательно, векторы $e_1, \ldots, e_m, e_{m+1}, \ldots, e_k$ линейно зависимы и по утверждению 1° заданные векторы линейно зависимы.

Следствие. Если все векторы линейно независимы, то и любая их часть линейно независима.

Доказательство. $\sqcap$ Часть векторов линейно зависима. Тогда по теореме все векторы линейно зависимы. $\sqcup$ ■

Следующая теорема даёт критерии линейной зависимости векторов.

Теорема 4.2.2 (Критерии линейной зависимости).

1° *Вектор* e_1 *линейно зависим тогда и только тогда, когда он нулевой.*

2° *Векторы* $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно зависимы тогда и только тогда, когда хотя бы один из них линейно выражается через остальные векторы, т.е.* $\exists j : e_j = \sum\limits_{i \neq j} \beta_i e_i$, $\beta_i \in P$.

3° *Векторы* $e_1, e_2, \ldots, e_k \in E\,(k > 1)$ *линейно зависимы тогда и только тогда, когда* $e_1 = \theta$ *или один из них есть линейная комбинация предыдущих, т.е.* $\exists j > 1$ $e_j = \sum\limits_{i=1}^{j-1} \beta_i e_i$, $\beta_i \in P$.

Доказательство. 1° Если вектор e_1 линейно зависим, то найдётся число $\alpha_1 \neq 0$ такое, что $\alpha_1 e_1 = \theta$. Отсюда $e_1 = \theta$.

Обратно. Если $e_1 = \theta$, то $1 e_1 = \theta$ — нетривиальная линейная комбинация равна θ. Следовательно, нулевой вектор линейно зависим.

2° Если векторы $e_1, e_2, \ldots, e_k$ линейно зависимы, то $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$ для некоторых чисел α_i, $i = \overline{1,m}$, среди которых есть $\alpha_j \neq 0$. Тогда $\alpha_j e_j + \sum\limits_{i \neq j} \alpha_i e_i = \theta \Rightarrow \alpha_j e_j = \sum\limits_{i \neq j} (-\alpha_i) e_i \Rightarrow e_j = \sum\limits_{i \neq j} (-\alpha_i/\alpha_j) e_i \Rightarrow e_j = \sum\limits_{i \neq j} \beta_i e_i$, где $\beta_i = -\alpha_i/\alpha_j \Rightarrow$ e_j выражается через остальные векторы.

Обратно. Пусть один из векторов $e_1, e_2, \ldots, e_k$ линейно выражается через остальные векторы: $e_j = \sum\limits_{i \neq j} \beta_i e_i$. Отсюда $1e_j + \sum\limits_{i \neq j} (-\beta_i) e_i = \theta$. Следовательно, векторы линейно зависимы.

3° Пусть векторы $e_1, e_2, \ldots, e_k \in E\,(k > 1)$ линейно зависимы.

При $e_1 = \theta$ утверждение верно.

При $e_1 \neq \theta$ из линейной зависимости векторов $e_1, e_2, \ldots, e_k$ следует, что $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$ для некоторых чисел α_i, $i = \overline{1,m}$, среди которых есть $\alpha_n \neq 0$. Обозначим через j наибольший среди индексов i ненулевых коэффициентов α_i. Если бы $j = 1$, то $\alpha_1 \neq 0$ и $\alpha_2 = \ldots = \alpha_k = 0$. Тогда $\sum\limits_{i=1}^{k} \alpha_i e_i = \alpha_1 e_1$ и $\alpha_1 e_1 = \theta$, где $\alpha_1 \neq 0$. Отсюда $e_1 = \theta$, что неверно. Следовательно, $j > 1$, $\alpha_j \neq 0$ и $\sum\limits_{i=1}^{j} \alpha_i e_i = \theta$. Из этого равенства выражаем вектор e_j : $\alpha_j e_j + \sum\limits_{i=1}^{j-1} \alpha_i e_i = \theta \Rightarrow \alpha_j e_j = \sum\limits_{i=1}^{j-1} (-\alpha_i e_i) \Rightarrow e_j = \sum\limits_{i=1}^{j-1} \left(\dfrac{-\alpha_i}{\alpha_j}\right) e_i$.

Обратно. Пусть $e_1 = \theta$ или вектор e_j линейно выражается через предыдущие векторы $e_j = \sum\limits_{i=1}^{j-1} \beta_i e_i$. Тогда e_1 линейно зависим или $e_1, \ldots, e_j$ линейно зависимы. Причём это часть всех векторов $e_1, \ldots, e_k$, поэтому все векторы линейно зависимы.

Следствия. (Критерии линейной независимости векторов).

1° *Вектор* e_1 *линейно независим* $\Leftrightarrow$ $e_1 \neq \theta$.

2° *Векторы* $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно независимы* $\Leftrightarrow$ *ни один* e_j, $j = \overline{1,k}$ *не является линейной комбинацией остальных векторов, т. е. при любом* j *выполнено неравенство* $e_j \neq \sum\limits_{i \neq j} \beta_i e_i$ *для любых* $\beta_i \in P$.

3° *Векторы* $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно независимы* $\Leftrightarrow$ $e_1 \neq \theta$, *и ни один* $e_j, j > 1$ *не является линейной комбинацией предыдущих векторов, т. е. при любом* $j > 1$ *выполнено неравенство* $e_j \neq \sum\limits_{i \neq j} \beta_i e_i$

для любых $\beta_i \in P$.

Справедливость следствий $1^{\circ}-3^{\circ}$ устанавливается доказательством «от противного» или следуют из математической логики: утверждения $P \Leftrightarrow Q$ и $\neg P \Leftrightarrow \neg Q$ эквивалентны. ■

Приведём ещё несколько условий линейной зависимости и линейной независимости векторов.

Теорема 4.2.3. 1° *Если векторы* $e_1, e_2, \ldots, e_k$ *линейно независимы, а векторы* $e_1, e_2, \ldots, e_k, x$ *линейно зависимы, то* x *линейно выражается через* $e_1, e_2, \ldots, e_k$.

2° 1) *Если существует вектор* x, *разложение которого по векторам* $e_1, e_2, \ldots, e_k$ *единственно, то* $e_1, e_2, \ldots, e_k$ *линейно независимы.*

2) *Если векторы* $e_1, e_2, \ldots, e_k$ *линейно независимы, то любой вектор* x, *являющийся линейной комбинацией этих векторов, имеет единственное разложение по ним.*

Доказательство. 1° Из линейной зависимости векторов $e_1, e_2, \ldots$ $\ldots, e_k, x$ следует, что первый вектор нулевой или какой-то вектор линейно выражается через предыдущие. Из линейной независимости векторов $e_1, e_2 \ldots, e_k$ следует, что первый вектор ненулевой и ни один из этих векторов линейно не выражается через предыдущие. Поэтому только вектор x линейно выражается через предыдущие векторы.

2° 1) Пусть вектор x имеет единственное разложение по векторам $e_1, e_2, \ldots, e_k$: $x = \sum\limits_{i=1}^{k} \beta_i e_i$, где $\beta_i \in P$. Требуется доказать, что векторы $e_1, e_2, \ldots, e_k$ линейно независимы.

$\sqcap$ Векторы $e_1, e_2 \ldots, e_k$ линейно зависимы. Тогда $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$ для некоторых коэффициентов $\alpha_1, \alpha_2, \ldots, \alpha_k$, среди которых есть $\alpha_j \neq 0$. Сложив равенства $x = \sum\limits_{i=1}^{k} \beta_i e_i$ и $\theta = \sum\limits_{i=1}^{k} \alpha_i e_i$, получим второе разложение вектора $x = \sum\limits_{i=1}^{k} (\beta_i + \alpha_i) e_i$. Разложения различны, т. к. в силу $\alpha_j \neq 0$ имеем $\beta_j \neq \beta_j + \alpha_j$. $\sqcup$ Следовательно, векторы $e_1, e_2, \ldots, e_k$ линейно независимы.

2) Каждый вектор x, который является линейной комбинацией данных векторов $e_1, e_2, \ldots, e_k$, имеет разложение $x = \sum\limits_{i=1}^{k} \alpha_i e_i$ по этим векторам. Докажем, что это разложение единственно при условии линейной независимости векторов $e_1, e_2, \ldots, e_k$.

$\sqcap$ Для вектора $x \in E$ существует два различных разложения $x = \sum\limits_{i=1}^{k} \alpha_i e_i$ и $x = \sum\limits_{i=1}^{k} \beta_i e_i$, где $\alpha_j \neq \beta_j$ для некоторого j. Вычитание двух разложений приводит к равенству $\sum\limits_{i=1}^{k} (\alpha_i - \beta_i) e_i = \theta$, где $\alpha_j - \beta_j \neq 0$ для некоторого j. Получена нетривиальная линейная комбинация векторов $e_1, e_2, \ldots, e_k$, равная θ, значит эти векторы линейно зависимы. $\sqcup$ ■

Задачи.

Задача 4.2.1. *Определите, являются ли линейно независимыми матрицы* $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & -1 \\ 4 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$.

Задача 4.2.2. *В пространстве многочленов* $R[x]$ *заданы многочлены* $f_1(x) = -1 + 2x - x^2 + 2x^3$, $f_2(x) = 2 + x - 2x^2 + x^3$, $f_3(x) = 3 + 4x - 5x^2 + 4x^3$, $f_4(x) = 4 - 3x - 3x^3$. *Проверьте, являются ли они линейно зависимыми. Если да, то укажите общий вид их нетривиальных линейных комбинаций, равных нулевому многочлену, и один частный пример такой линейной комбинации.*

Решение задач.

4.2.1 Составим линейную комбинацию матриц и приравняем её нулевой матрице θ :

$$\alpha_1 \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 & -1 \\ 4 & 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

После выполнения линейных операций умножения матриц на числа и последующего сложения матриц имеем равенство

$$\begin{pmatrix} \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 & 2\alpha_1 - \alpha_2 - \alpha_3 - 2\alpha_4 \\ 2\alpha_2 + 4\alpha_3 & \alpha_1 - \alpha_2 - \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Приравняв соответствующие элементы матриц в левой и правой части равенства, получим систему линейных уравнений

$$\begin{cases} \alpha_1+\alpha_2+\alpha_3+\alpha_4=0, \\ 2\alpha_1-\alpha_2-\alpha_3-2\alpha_4=0, \\ 2\alpha_2+4\alpha_3=0, \\ \alpha_1-\alpha_2-\alpha_4=0. \end{cases}$$

Запишем и преобразуем расширенную матрицу системы:

$$\begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{array} \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 2 & -1 & -1 & -2 & 0 \\ 0 & 2 & 4 & 0 & 0 \\ 1 & -1 & 0 & -1 & 0 \end{array}\right) \\ r_2 := r_2 - 2r_1, \\ r_4 := r_4 - r_1, \end{array} \Leftrightarrow$$

строки обозначены r_i, в r_i записаны коэффициенты при неизвестных $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ и свободный член i-го уравнения; делаем линейные преобразования строк так, чтобы матрица системы (до черты) приобрела верхнюю ступенчатую форму; запись $r_2 := r_2 - 2r_1$ означает, что из 2-й строки вычитается удвоенная 1-я строка и результат записывается во 2-ю строку;

$$\Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & -4 & 0 \\ 0 & 2 & 4 & 0 & 0 \\ 0 & -2 & -1 & -2 & 0 \end{array}\right) \\ r_3 := r_3 + 2r_2, \\ r_4 := r_4 - 2r_2, \end{array} \Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & -4 & 0 \\ 0 & 0 & 6 & -8 & 0 \\ 0 & 0 & -3 & 6 & 0 \end{array}\right) \\ r_2 := -r_2, r_4 := r_4 + r_3/2, \\ r_3 := r_3/2, \end{array} \Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 4 & 0 \\ 0 & 0 & 3 & -4 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{array}\right) \end{array}.$$

Из эквивалентной системы с преобразованной матрицей $\begin{cases} \alpha_1+\alpha_2+\alpha_3+\alpha_4=0, \\ \alpha_2-\alpha_3+4\alpha_4=0, \\ 3\alpha_3-4\alpha_4=0, \\ 2\alpha_4=0 \end{cases}$ обратным ходом находим $\begin{cases} \alpha_1=0, \\ \alpha_2=0, \\ \alpha_3=0, \\ \alpha_4=0. \end{cases}$

Это означает, что только тривиальная линейная комбинация матриц равна нулевой матрице и данные матрицы линейно независимы. ■

4.2.2 Составим линейную комбинацию многочленов и приравняем её нулевому многочлену:

$$\alpha_1 f_1(x)+\alpha_2 f_2(x)+\alpha_3 f_3(x)+\alpha_4 f_4(x)=0 \Leftrightarrow$$

$\alpha_1(-1+2x-x^2+2x^3)+\alpha_2(2+x-2x^2+x^3)+\alpha_3(3+4x-5x^2+4x^3)+$ $+\alpha_4(4-3x-3x^3)=0 \Leftrightarrow (-\alpha_1+2\alpha_2+3\alpha_3+4\alpha_4)+(2\alpha_1+\alpha_2+4\alpha_3-3\alpha_4)x+$ $+(-\alpha_1-2\alpha_2-5\alpha_3)x^2+(2\alpha_1+\alpha_2+4\alpha_3-3\alpha_4)x^3=0 \Leftrightarrow$ {многочлен равен нулевому многочлену, только если все его коэффициенты равны нулю}

$$\begin{cases} -\alpha_1+2\alpha_2+3\alpha_3+4\alpha_4=0, \\ 2\alpha_1+\alpha_2+4\alpha_3-3\alpha_4=0, \\ -\alpha_1-2\alpha_2-5\alpha_3=0, \\ 2\alpha_1+\alpha_2+4\alpha_3-3\alpha_4=0, \end{cases}$$

запишем и преобразуем расширенную матрицу системы $\Leftrightarrow$

$$\begin{array}{c} \\ r_1 \\ r_2 \\ r_3 \\ r_4 \end{array} \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} -1 & 2 & 3 & 4 & 0 \\ 2 & 1 & 4 & -3 & 0 \\ -1 & -2 & -5 & 0 & 0 \\ 2 & 1 & 4 & -3 & 0 \end{array}\right) \end{array} \Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} -1 & 2 & 3 & 4 & 0 \\ 0 & 5 & 10 & 5 & 0 \\ 0 & -4 & -8 & -4 & 0 \\ 0 & 5 & 10 & 5 & 0 \end{array}\right) \end{array} \Leftrightarrow$$

$$\begin{array}{c} r_2 := r_2 + 2r_1, \\ r_3 := r_3 - r_1, \\ r_4 := r_4 + 2r_1 \end{array} \qquad \begin{array}{c} r_3 := r_3 + 4r_2/5, \\ r_4 := r_4 - r_2, \\ r_2 := r_2/5, r_1 := -r_1, \end{array}$$

$$\Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & -2 & -3 & -4 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array} \Leftrightarrow \begin{array}{c} \begin{array}{ccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \text{с.ч} \end{array} \\ \left(\begin{array}{cccc|c} 1 & 0 & 1 & -2 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array} \Leftrightarrow$$

$$r_1 := r_1 + 2r_2,$$

Здесь первые ненулевые коэффициенты в строках стоят при α_1 и α_2, поэтому в преобразованной системе выразим α_1 и α_2 только через α_3 и α_4.

$$\Leftrightarrow \begin{cases} \alpha_1 + \alpha_3 - 2\alpha_4 = 0, \\ \alpha_2 + 2\alpha_3 + \alpha_4 = 0, \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = -\alpha_3 + 2\alpha_4, \\ \alpha_2 = -2\alpha_3 - \alpha_4, \end{cases}$$

где α_3, α_4 принимают любые значения.

Отсюда следует, что

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \alpha_4 f_4(x) = 0 \Leftrightarrow \begin{cases} \alpha_1 = -\alpha_3 + 2\alpha_4, \\ \alpha_2 = -2\alpha_3 - \alpha_4, \end{cases} \forall \alpha_3, \alpha_4 \in R.$$

Для исключения тривиальной линейной комбинации нужно взять любые значения $\alpha_3, \alpha_4 \in R$, кроме $\alpha_3 = \alpha_4 = 0$. Следовательно, общий вид нетривиальных линейных комбинаций:

$$(-\alpha_3 + 2\alpha_4) f_1(x) + (-2\alpha_3 - \alpha_4) f_2(x) + \alpha_3 f_3(x) + \alpha_4 f_4(x) = 0,$$

при любых значениях α_3, $\alpha_4 \in R$, кроме $\alpha_3 = \alpha_4 = 0$.

Частный пример получим при $\alpha_3 = 1$, $\alpha_4 = 0$, тогда $\alpha_1 = -\alpha_3 + 2\alpha_4 = -1$, $\alpha_2 = -2\alpha_3 - \alpha_4 = -2$, и линейная комбинация имеет вид:

$$-f_1(x) - 2f_2(x) + f_3(x) + 0f_4(x) = 0 \text{ или } f_1(x) + 2f_2(x) - f_3(x) = 0. \blacksquare$$

4.3. База и ранг системы векторов. Критерий базы

Будем рассматривать конечные системы векторов линейного пространства над полем P.

Определение 1. 1) *Базой системы векторов называется её линейно независимая подсистема (часть), через которую линейно выражается любой вектор системы.*

2) *В данной системе векторов линейно независимая подсистема из r векторов называется максимальной линейно независимой подсисте-*

мой, если в этой системе нет линейно независимых подсистем более чем из r векторов.

3) *Рангом система векторов $e_1, e_2, \ldots, e_k$ называется число векторов её базы. Обозначение: $rg(e_1, e_2, \ldots, e_k)$.*

Теорема 4.3.1 (Критерий базы). *Подсистема системы векторов образует базу тогда и только тогда, когда является максимальной линейно независимой подсистемой.*

Доказательство.

Будем использовать понятие ранга матрицы, теорему о базисном миноре и её следствие.

Определение 2. Рангом *ненулевой матрицы называется максимальный порядок ненулевых миноров. Ранг нулевой матрицы считается равным 0. Ранг матрицы $\mathcal{A}$ обозначается $r_{\mathcal{A}}$ или $rg(A)$.*

В ненулевой матрице любой ненулевой минор порядка, равного рангу, называется **базисным минором**. *Строки и столбцы матрицы, на которых расположен базисный минор, называются базисными.* ■

Теорема 4.3.2 (Теорема о базисном миноре).

В ненулевой матрице базисные столбцы (строки) линейно независимы и любой столбец (любая строка) линейно выражается через базисные столбцы (строки).

Следствие из теоремы о базисном миноре.
Ранг матрицы равен максимальному числу линейно независимых столбцов (строк). ■

Продолжим доказательство теоремы 4.3.1.

I. Пусть система из k векторов $e_1, e_2, \ldots, e_k$ содержит базу из r линейно независимых векторов.

а) При $r = k$ база совпадает со всей системой и является максимальной линейно независимой подсистемой.

б) При $r < k$ покажем, что любые $r + 1$ векторов системы линейно зависимы и тогда база является максимальной линейно независимой подсистемой.

Будем считать, что в системе векторов первые r векторов $e_1, \ldots, e_r$ образуют базу. Из определения 1 следует, что каждый вектор системы линейно выражается через векторы базы: $e_j = a_{1j}e_1 + a_{2j}e_2 + \ldots + a_{rj}e_r = \sum\limits_{i=1}^{r} a_{ij}e_i$, $j = \overline{1,k}$ где $a_{ij} \in P$ — числа. Составим $r \times k$-матрицу $A = (a_{ij})$, $i = \overline{1,r}$, $j = \overline{1,k}$. В ней j-й столбец состоит из коэффициентов a_{ij}, $i = \overline{1,r}$ разложения вектора e_j по векторам базы. Поскольку разложение вектора e_j при $j \leqslant r$ имеет вид $e_j = 1e_j = 0e_1 + \ldots + 0e_{j-1} + 1e_j + 0e_{j+1} + \ldots + 0e_r$, то $a_{ij} = 0$ для $i \neq j$ и $a_{ij} = 1$ для $i = j$.

Изобразим вид матрицы: $A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & a_{1,r+1} & \ldots & a_{1k} \\ 0 & 1 & 0 & \ldots & 0 & a_{2,r+1} & \ldots & a_{2k} \\ 0 & 0 & 1 & \ldots & 0 & a_{3,r+1} & \ldots & a_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & a_{r,r+1} & \ldots & a_{rk} \end{pmatrix}$.

Здесь первые r столбцов образуют единичную подматрицу, определитель которой равен 1. Поэтому ранг матрицы A равен r. Из следствия теоремы о базисном миноре следует, что максимальное число линейно независимых столбцов матрицы A равно r. Тогда произвольные $r+1$ столбцов матрицы A линейно зависимы. Докажем, что соответствующие этим столбцам векторы линейно зависимы. Пусть номера этих столбцов образуют множество M. Поскольку эти столбцы линейно зависимы, то существуют коэффициенты β_j, $j \in M$, не все равные 0 и такие, что

$$\sum_{j \in M} \beta_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{rj} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} \sum\limits_{j \in M} \beta_j a_{1j} \\ \sum\limits_{j \in M} \beta_j a_{2j} \\ \vdots \\ \sum\limits_{j \in M} \beta_j a_{rj} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow \sum_{j \in M} \beta_j a_{ij} = 0,$$

$i = \overline{1,r}$.

Рассмотрим и преобразуем линейную комбинацию векторов e_j, $j \in M$ с выбранными выше коэффициентами β_j :

$$\sum_{j \in M} \beta_j e_j = \sum_{j \in M} \beta_j \sum_{i=1}^{r} a_{ij}e_i = \sum_{j \in M} \sum_{i=1}^{r} \beta_j a_{ij} e_i = \sum_{i=1}^{r} \sum_{j \in M} \beta_j a_{ij} e_i =$$

$= \sum\limits_{i=1}^{r} \Big(\sum\limits_{j \in M} \beta_j a_{ij} \Big) e_i = \sum\limits_{i=1}^{r} 0 e_i = \sum\limits_{i=1}^{r} \theta = \theta$. Нетривиальная линейная комбинация $\sum\limits_{j \in M} \beta_j e_j$ равна θ, поэтому $r+1$ векторов e_j, $j \in M$ — линейно зависимы.

Из а) и б) следует, что база системы векторов является максимальной линейно независимой подсистемой этой системы.

II. Переставляя векторы, если нужно, считаем, что r первых векторов $e_1, \ldots, e_r$ образуют максимальную линейно независимую подсистему в системе векторов $e_1, \ldots, e_r, e_{r+1}, \ldots, e_k$.

Покажем, что любой вектор e_j системы линейно выражается через векторы подсистемы $e_1, \ldots, e_r$.

При $j \leqslant r$ это очевидно: $e_j = e_j$.

При $j > r$ рассмотрим подсистему из $r+1$ векторов $e_1, \ldots, e_r, e_j$. Она линейно зависима, т.к. любая линейно независимая подсистема содержит не более r векторов. Кроме того, векторы $e_1, \ldots, e_r$ линейно независимы. Тогда из утверждения 1° теоремы 4.2.3 следует, что e_j линейно выражается через векторы $e_1, \ldots, e_r$.

Следовательно, подсистема $e_1, \ldots, e_r$ образует базу исходной системы $e_1, \ldots, e_r, e_{r+1}, \ldots, e_k$.

Следствие. *Все базы данной системы векторов содержат одинаковое число векторов, равное максимальному числу линейно независимых векторов системы.*

Доказательство.

Предположим, что $B_1 = \{e_1, \ldots, e_r\}$ и $B_2 = \{f_1, \ldots, f_m\}$ — две базы в данной системе. По теореме база является максимальной линейно независимой подсистемой. Тогда $r \geqslant m$, т. к. B_1 — база, $m \geqslant r$, т. к. B_2 — база, и $r = m$. ■

Задачи.

Задача 4.3.1. *Докажите следующие утверждения.*

а) Векторы, содержащие нулевой вектор, линейно зависимы.

б) Векторы линейно независимы тогда и только тогда, когда любой вектор, являющийся линейной комбинацией этих векторов, имеет

единственное разложение по ним.

в) Ненулевые векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно зависимы тогда и только тогда, когда один из них линейно выражается через предыдущие, т.е.* $\exists j > 1: \ e_j = \sum\limits_{i=1}^{j-1} \beta_i e_i$.

г) Ненулевые векторы $e_1, e_2, \ldots, e_k$, $k > 1$ *линейно независимы тогда и только тогда, когда ни один из них линейно не выражается через предыдущие, т.е.* $\forall j > 1: \ e_j \neq \sum\limits_{i=1}^{j-1} \beta_i e_i$ *при любых числах* $\beta_1, \ldots, \beta_{j-1}$.

Задача 4.3.2. *В пространстве* R^3 *заданы векторы* $e_1 = (2, 3, 4)$, $e_2 = (3, 4, 5), e_3 = (4, 5, 6)$, $e_4 = (5, 7, 9)$. *Найдите базу этой системы векторов и выразите все векторы системы через векторы базы.*

Решение задач.

4.3.2 Найдём базу (максимальную линейно независимую подсистему) в системе векторов e_1, e_2, e_3, e_4. Для этого составим линейную комбинацию этих векторов с коэффициентами x_1, x_2, x_3, x_4 и приравняем её нулевому вектору:

$$x_1 e_1 + x_2 e_2 + x_3 e_3 + x_4 e_4 = \theta.$$

Отметим, что линейная зависимость координатных строк эквивалентна линейной зависимости координатных столбцов. Поэтому запишем это равенство для координатных столбцов векторов

$$x_1 \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} + x_3 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + x_4 \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (65)$$

Уравнение (65) эквивалентно системе линейных уравнений

$$\begin{cases} 2\,x_1 + 3\,x_2 + 4\,x_3 + 5\,x_4 = 0, \\ 3\,x_1 + 4\,x_2 + 5\,x_3 + 7\,x_4 = 0, \\ 4\,x_1 + 5\,x_2 + 6\,x_3 + 9\,x_4 = 0. \end{cases} \qquad (66)$$

Преобразуем расширенную матрицу системы к верхней ступенчатой форме с помощью элементарных преобразований строк r_1, r_2, r_3 :

$$\begin{array}{c} r_1 \\ r_2 \\ r_3 \\ r_2 := 2r_2, r_3 := r_3 - 2r_1 \end{array} \left(\begin{array}{cccc|c} 2 & 3 & 4 & 5 & 0 \\ 3 & 4 & 5 & 7 & 0 \\ 4 & 5 & 6 & 9 & 0 \end{array}\right) \Leftrightarrow \begin{array}{c} \\ \\ \\ r_2 := r_2 - 3r_1 \end{array} \left(\begin{array}{cccc|c} 2 & 3 & 4 & 5 & 0 \\ 6 & 8 & 10 & 14 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{array}\right) \Leftrightarrow \begin{array}{c} \\ \\ \\ r_3 := r_3 - r_2, r_2 := -r_2 \end{array} \left(\begin{array}{cccc|c} 2 & 3 & 4 & 5 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{array}\right) \Leftrightarrow$$

$$\Leftrightarrow \left(\begin{array}{cccc|c} 2 & 3 & 4 & 5 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{cccc|c} 2 & 0 & -2 & 2 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

$r_1 := r_1 - 3r_2$ $\qquad$ $r_1 := r_1/2$

Ранги всех матриц равны, т. к. ранг сохраняется при элементарных преобразованиях. Ранг последней матрицы равен 2, т.к. минор $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$ является её ненулевым минором порядка 2 и нет ненулевых миноров большего порядка. Этот минор является базисным минором последней матрицы. Базисным минором исходной матрицы является минор $\begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -1$, расположенный в первых двух столбцах и строках. По теореме о базисном миноре базисные столбцы матрицы линейно независимы и все столбцы линейно выражаются через базисные столбцы. Поэтому базой системы векторов e_1, e_2, e_3, e_4 являются первые два вектора e_1, e_2.

Для выражения всех векторов через базисные векторы найдём решение системы уравнений (66). Возьмём эквивалентную систему с последней преобразованной матрицей $\begin{cases} x_1 - x_3 + x_4 = 0, \\ x_2 + 2\,x_3 + x_4 = 0 \end{cases}$ и запишем общее решение в виде выражения базисных неизвестных через свободные:

$$\begin{cases} x_1 = x_3 - x_4, \\ x_2 = -2\,x_3 - x_4, \end{cases} \quad \forall x_3,\, x_4.$$

Выразим все векторы e_j через векторы базы.

При $j \leqslant 2$ это очевидно: $e_1 = e_1$, $e_2 = e_2$

При $j > 3$ воспользуемся равенством

Отсюда после подстановки $\begin{cases} x_1 = x_3 - x_4, \\ x_2 = -2\,x_3 - x_4 \end{cases}$ получим

$$(x_3 - x_4)e_1 + (-2\,x_3 - x_4)e_2 + x_3e_3 + x_4e_4 = \theta.$$

При $x_3 = 1$, $x_4 = 0$ находим $e_1 - 2e_2 + e_3 = \theta \Rightarrow e_3 = -e_1 + 2e_2$.

При $x_3 = 0$, $x_4 = 1$ находим $-e_1 - e_2 + e_4 = \theta \Rightarrow e_4 = e_1 + e_2$. ■

4.4. Базис линейного пространства. Координаты вектора в базисе и их свойства

В этом разделе вводятся понятия базиса и координат вектора в линейном пространство E над полем P.

Определение 1. *Назовём базисом линейного пространства E упорядоченный набор векторов $e_1, e_2, \ldots, e_n \in E$ таких, что:*

1) *$e_1, e_2, \ldots, e_n$ линейно независимы;*

2) *любой вектор $x \in E$ линейно выражается через $e_1, e_2, \ldots, e_n$, т. е. найдутся числа $\alpha_1, \alpha_2, \ldots, \alpha_n \in P$ такие, что*

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = \sum_{i=1}^{n} \alpha_i e_i. \tag{67}$$

при этом числа $\alpha_1, \alpha_2, \ldots, \alpha_n$ называются координатами вектора x в базисе (или относительно базиса) $e = (e_1, e_2, \ldots, e_n)$.

Обозначим координатный столбец вектора x так: $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$,

тогда равенство (67) можно записать в виде

$$x = (e_1\ e_2\ \ldots\ e_n) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} \quad \Leftrightarrow \quad x = e[x]_e, \tag{68}$$

где в правой части равенства стоит произведение матриц.

Единственность координат вектора в базисе следует из теоремы 4.2.3.

Отметим, что для нулевого вектора $x = \theta$ равенство (67) выполнено для $\alpha_i = 0$, $i = \overline{1, n}$, т.е. нулевой вектор имеет нулевые координаты в любом базисе: $[\theta]_e = \begin{pmatrix} 0, 0, \ldots, 0 \end{pmatrix}^\top$.

Для j-го вектора базиса $x = e_j$ равенство (67) выполнено для $\alpha_i = 0$, $i \neq j$ и $\alpha_j = 1$, т.е. все координаты j-го вектора базиса относительно

этого базиса равны 0, кроме j-й координаты, которая равна 1: $[e_j]_e = \big(0, \ldots, \overset{j}{1}, \ldots, 0\big)^\top$.

Теорема 4.4.1 (Свойства координат).

1° *Координаты любого вектора в фиксированном базисе определяются однозначно.*

2° *При сложении векторов складываются их соответствующие координаты или каждая координата суммы векторов равна сумме их соответствующих координат:*

$$\forall f_1,\, f_2,\, \ldots,\, f_k \in E \quad \left[\sum_{j=1}^{k} f_j\right]_e = \sum_{j=1}^{k} [f_j]_e.$$

3° *При умножении вектора на число все его координаты умножаются на это число:*

$$\forall f \in E,\, \beta \in P \quad [\beta f]_e = \beta [f]_e.$$

4° *Каждая координата линейной комбинации векторов равна линейной комбинации соответствующих координат этих векторов с теми же коэффициентами:*

$$\forall f_1,\, f_2,\, \ldots,\, f_k \in E,\, \beta_1,\, \beta_2,\, \ldots,\, \beta_k \in P \quad \left[\sum_{j=1}^{k} \beta_j f_j\right]_e = \sum_{j=1}^{k} \beta_j [f_j]_e.$$

5° *Векторы линейно зависимы (независимы) $\Leftrightarrow$ их координатные столбцы линейно зависимы (независимы).*[3]

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ — базис линейного пространства E.

1° Если $x = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $x = \sum\limits_{i=1}^{n} \alpha'_i e_i$, то $x - x = \sum\limits_{i=1}^{n} \alpha_i e_i - \sum\limits_{i=1}^{n} \alpha'_i e_i$ и тогда $\theta = \sum\limits_{i=1}^{n} (\alpha_i - \alpha'_i) e_i$. В силу линейной независимости векторов базиса отсюда следует, что все коэффициенты $(\alpha_i - \alpha'_i) = 0$, т.е. $\alpha_i = \alpha'_i$ для всех i.

2° Применим метод математической индукции.

1) Докажем утверждение при $k = 2$.

[3] Мы рассматриваем координатные столбцы, состоящие из n чисел из поля P, как элементы P^n или $P^{n\times 1}$ — линейного пространства матриц-столбцов с операциями сложения матриц и умножения матриц на числа из поля P.

Пусть заданы векторы $f_j = \sum\limits_{i=1}^{n} \alpha_{ij} e_i$ с координатами $\alpha_{1j}, \alpha_{2j}, \ldots,$ α_{nj}, $j = \overline{1,2}$. Тогда $f_1 + f_2 = \sum\limits_{i=1}^{n} \alpha_{i1} e_i + \sum\limits_{i=1}^{n} \alpha_{i2} e_i = \sum\limits_{i=1}^{n} (\alpha_{i1} + \alpha_{i2}) e_i$. Имеем i-е координаты f_1, f_2, $f_1 + f_2$ — это соответственно α_{i1}, α_{i2}, $\alpha_{i1} + \alpha_{i2}$, т.е. при сложении двух векторов складываются их соответствующие координаты. Этот результат можно изобразить в виде координатных столбцов.

Координатные столбцы векторов f_1, f_2, $f_1 + f_2$ соответственно равны:

$$[f_1]_e = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{pmatrix}, \quad [f_2]_e = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{n2} \end{pmatrix}, \quad [f_1 + f_2]_e = \begin{pmatrix} \alpha_{11} + \alpha_{12} \\ \alpha_{21} + \alpha_{22} \\ \vdots \\ \alpha_{n1} + \alpha_{n2} \end{pmatrix}.$$

Тогда $[f_1 + f_2]_e = \begin{pmatrix} \alpha_{11} + \alpha_{12} \\ \alpha_{21} + \alpha_{22} \\ \vdots \\ \alpha_{n1} + \alpha_{n2} \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{pmatrix} + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{n2} \end{pmatrix} = [f_1]_e + [f_2]_e,$

поэтому $[f_1 + f_2]_e = [f_1]_e + [f_2]_e$, то есть утверждение верно при $k = 2$.

2) Предположим, что утверждение верно при $k = m$, и докажем, что утверждение верно при $k = m + 1$.

Пусть заданы векторы f_j, $j = \overline{1, m+1}$. Запишем доказательство с помощью координатных столбцов.

Сумму $m+1$ векторов f_j, $j = \overline{1, m+1}$ запишем в виде суммы двух векторов $\sum\limits_{j=1}^{m+1} f_j = \sum\limits_{j=1}^{m} f_j + f_{m+1}$. Приравняем координатные столбцы равных векторов $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \left[\sum\limits_{j=1}^{m} f_j + f_{m+1}\right]_e$, и применим утверждение при $k = 2$. Получим $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \left[\sum\limits_{j=1}^{m} f_j\right]_e + \left[f_{m+1}\right]_e$. Далее, используя предположение индукции $\left[\sum\limits_{j=1}^{m} f_j\right]_e = \sum\limits_{j=1}^{m} [f_j]_e$, запишем $\left[\sum\limits_{j=1}^{m+1} f_j\right]_e = \sum\limits_{j=1}^{m} [f_j]_e +$ $+ [f_{m+1}]_e$. Отсюда следует справедливость утверждения для $k = m + 1$:

$$\left[\sum_{j=1}^{m+1} f_j\right]_e = \sum_{j=1}^{m+1} [f_j]_e .$$

3° Пусть заданы вектор $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ с координатами $\alpha_1, \alpha_2, \ldots, \alpha_n$ и число $\beta \in P$. Тогда $\beta f = \beta \sum\limits_{i=1}^{n} \alpha_i e_i = \sum\limits_{i=1}^{n} (\beta\alpha_i) e_i$. Имеем i-е координаты f и βf — это соответственно α_i и $\beta\alpha_i$, т.е. при умножении вектора на число все его координаты умножаются на это число. Этот результат можно изобразить в виде координатных столбцов: $[\beta f]_e = \beta [f]_e$.

4° Пусть заданы векторы $f_j \in E$, $j = \overline{1,k}$ и их линейная комбинация $\sum\limits_{j=1}^{k} \beta_j f_j$ с коэффициентами $\beta_j \in P$, $j = \overline{1,k}$. Запишем координатный столбец линейной комбинации: $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e$. К этому выражению применим сначала свойство 2°, а затем свойство 3° : $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = \sum\limits_{j=1}^{k} [\beta_j f_j]_e =$ $= \sum\limits_{j=1}^{k} \beta_j [f_j]_e$. Следовательно, $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = \sum\limits_{j=1}^{k} \beta_j [f_j]_e$.

5° Докажем утверждение о линейной зависимости:
векторы $f_j \in E$, $j = \overline{1,k}$ линейно зависимы $\Leftrightarrow$ координатные столбцы векторов $[f_j]_e$, $j = \overline{1,k}$ линейно зависимы.

Пусть векторы f_j, $j = \overline{1,k}$ линейно зависимы. Тогда существует нетривиальная линейная комбинация векторов, равная θ : $\sum\limits_{j=1}^{k} \beta_j f_j = \theta$, где β_j, $j = \overline{1,k}$ — числа, среди которых есть $\beta_m \neq 0$. Равенство векторов $\sum\limits_{j=1}^{k} \beta_j f_j = \theta$ эквивалентно равенству координатных столбцов векторов $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = [\theta]_e$. В силу утверждения 4°, последнее равенство запишем в виде $\sum\limits_{j=1}^{k} \beta_j [f_j]_e = [\theta]_e$. Так как $[\theta]_e = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, мы получили нетривиальную линейную комбинацию координатных столбцов векторов, равную нулевому столбцу. Следовательно, координатные столбцы векто-

ров $[f_j]_e$, $j = \overline{1,k}$ линейно зависимы.

Далее проводим рассуждения последнего абзаца в обратном порядке и получим доказательство обратного утверждения.

Пусть координатные столбцы векторов $[f_j]_e$, $j = \overline{1,k}$ линейно зависимы. Тогда найдётся нетривиальная линейная комбинация координатных столбцов $[f_j]_e$, $j = \overline{1,k}$ равная нулевому столбцу $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = [\theta]_e$: $\sum\limits_{j=1}^{k} \beta_j [f_j]_e = [\theta]_e$. где β_j, $j = \overline{1,k}$ — числа, среди которых есть $\beta_m \neq 0$. В силу утверждения 4°, запишем линейную комбинацию так: $\left[\sum\limits_{j=1}^{k} \beta_j f_j\right]_e = [\theta]_e$. Равенство координатных столбцов эквивалентно равенству векторов: $\sum\limits_{j=1}^{k} \beta_j f_j = \theta$. Получена нетривиальная линейная комбинация векторов f_j, $j = \overline{1,k}$, равная θ, и тогда векторы f_j, $j = \overline{1,k}$ линейно зависимы. ■

4.5. Размерность линейного пространства. Связь понятий базиса и размерности. Дополнение линейно независимой системы векторов до базиса

В этом разделе введём понятие размерности линейного пространство E над полем P и выясним связь понятий базиса и размерности.

Определение 1. *Размерностью пространства E назовём максимальное число линейно независимых векторов в E. Обозначение:* $\dim E$ *или* $\dim(E)$.

1) *Размерность нулевого пространства $E = \theta$ считается равной 0 (в нём нет линейно независимых векторов).*

2) *Число $n \in N$ является размерностью ненулевого пространства, если в нём найдутся n линейно независимых векторов, а любые $n+1$ векторов линейно зависимы.*

3) *Размерность считается равной $+\infty$, если в E для любого $n \in N$ найдутся n линейно независимых векторов.*

Говорят, что в случаях 1) *и* 2) E *— n-мерное пространство и* E *— конечномерное пространство, в случае* 3) E *— бесконечномерное пространство.*

Выясним связь понятий базиса и размерности линейного пространства.

Теорема 4.5.1. 1° *Если размерность линейного пространства* E *равна* $n\,(n \in N)$, *то в* E *существует базис из* n *векторов и любые* n *линейно независимых векторов из* E *образуют в нём базис.*

2° *Если некоторый базис линейного пространства* E *состоит из* n *векторов, то* $\dim(E) = n$.

Доказательство. 1° Пусть $\dim(E) = n$. Из определения размерности следует, что в E существуют n линейно независимых векторов. Докажем, что n произвольных линейно независимых векторов $e_1, e_2, \ldots, e_n$ образуют базис в E. Добавим к ним любой вектор $x \in E$. Из определению размерности следует, что $n+1$ векторов $e_1, e_2, \ldots, e_n, x$ линейно зависимы. В силу теоремы 4.2.3 вектор x линейно выражается через $e_1, e_2, \ldots, e_n$. Следовательно, $e_1, e_2, \ldots, e_n$ — базис в E.

2° Пусть $e = (e_1, e_2, \ldots, e_n)$ — базис в E. Тогда эти n векторов линейно независимы и, следовательно, в E существуют n линейно независимых векторов.

Докажем, что произвольные $n+1$ векторов $f_1, f_2, \ldots, f_{n+1} \in E$ линейно зависимы. Разложим каждый вектор f_j, $j = \overline{1, n+1}$ по базису e : $f_j = \sum\limits_{i=1}^{n} a_{ij} e_i$ и из координатных столбцов $[f_j]_e = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ образуем $n \times (n+1)$-матрицу $A = (a_{ij})$. Ранг $r_A \leqslant n < n+1$. Из следствия теоремы о базисном миноре следует, что столбцы матрицы A линейно зависимы, т. е. координатные столбцы $[f_j]_e$, $j = \overline{1, n+1}$ линейно зависимы и по теореме 4.4.1 векторы $f_1, f_2, \ldots, f_{n+1}$ линейно зависимы.

Итак, в E существуют n линейно независимых векторов, а любые $n+1$ векторов линейно зависимы, тогда $\dim(E) = n$. ■

Следствия. 1. *Если размерность линейного пространства E равна n $(n \in N)$, то любой базис в E состоит из n векторов.*

2. $\dim(P^n) = n$.

Доказательство. 1. $\sqcap$ Пусть некоторый базис в E содержит k векторов и $k \neq n$. Тогда из утверждения 2° теоремы 4.5.1 следует, что $\dim(E) = k$ и $k \neq n$, а это противоречит условию $\dim(E) = n$. $\sqcup$ Следовательно, любой базис в E состоит из n векторов.

2. Базис в P^n образуют n векторов $e_j = (0, \ldots, \overset{j}{1}, \ldots, 0)^\top$, $j = \overline{1, n}$, т. к. произвольный вектор $x = (x_1, x_2, \ldots, x_n)^\top \in P^n$ имеет единственное разложение $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$. ■

Теорема 4.5.2 (О дополнении до базиса). *Если размерность линейного пространства E равна $n\,(n \in N)$, то любые $k\,(k < n)$ линейно независимых векторов из E можно дополнить до базиса.*

Доказательство.

Пусть $e_1, \ldots, e_k\,(k < n)$ — данные линейно независимые векторы. Если бы любой вектор линейно выражался через данные векторы, то это был бы базис и $\dim(E) = k$, что противоречит условию. Значит, существует вектор e_{k+1}, который линейно не выражается через данные векторы. Векторы $e_1, e_2, \ldots, e_k, e_{k+1}$ линейно независимы, т. к. $e_1 \neq \theta$ и ни один из векторов линейно не выражается через предыдущие. Если $k + 1 = n$, то эти n векторов образуют базис, т. к. $\dim(E) = n$. При $k + 1 < n$ аналогично построим линейно независимую систему из $k + 2$ векторов и т. д. За $n - k$ таких шагов получим линейно независимую систему $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ из n векторов, которая в силу теоремы 4.5.1 образует базис. ■

Задача 4.5.1. *Найдите размерности пространств и линейных подпространств из задач 4.1.1–4.1.2*

4.6. Матрица перехода от одного базиса к другому. Связь координат вектора в различных базисах

В этом разделе рассмотрим различные базисы в n-мерном линейном пространстве E над полем P и выясним связь координат вектора в различных базисах.

Пусть в n-мерном линейном пространстве E над полем P заданы:

$$e = (e_1,\, e_2,\, \ldots,\, e_n) \text{ — «старый» базис,} \tag{69}$$

$$e' = (e'_1,\, e'_2,\, \ldots,\, e'_n) \text{ — «новый» базис.} \tag{70}$$

Разложим векторы нового базиса e' по векторам старого базиса e :

$$\begin{aligned} e'_1 &= t_{11}e_1 + t_{21}e_2 + \ldots + t_{n1}e_n, \\ e'_2 &= t_{12}e_1 + t_{22}e_2 + \ldots + t_{n2}e_n, \\ &\ldots \\ e'_n &= t_{1n}e_1 + t_{2n}e_2 + \ldots + t_{nn}e_n, \end{aligned} \tag{71}$$

или в общем виде $e'_j = \sum\limits_{i=1}^{n} t_{ij}e_i,\ j = \overline{1,n}.$

Из координатных столбцов $[e'_j]_e = \begin{pmatrix} t_{1j} \\ t_{2j} \\ \ldots \\ t_{nj} \end{pmatrix}$, $j = \overline{1,n}$ составим матрицу

$$\begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix}. \tag{72}$$

Определение 1. *Матрица, состоящая из координатных столбцов векторов базиса e' в базисе e, называется матрицей перехода от базиса e к базису e'. Обозначение: T или $T_{e\to e'}$.*

Замечание. Формулы (71) могут быть записаны в виде произведения

$$(e'_1,\, e'_2,\, \ldots,\, e'_n) = (e_1,\, e_2,\, \ldots,\, e_n) \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix} \tag{73}$$

$$\text{или } e' = eT_{e\to e'}.$$

Сформулируем и докажем свойства матриц перехода.

Теорема 4.6.1 (Свойства матриц перехода).

1° *Матрица перехода невырождена.*

2° *Если T матрица перехода от базиса e к базису e', то обратная матрица T^{-1} является матрицей перехода от базиса e' к базису e, т.е. $T^{-1}_{e\to e'} = T_{e'\to e}$.*

3° *Для любой невырожденной матрицы $T \in P^{n\times n}$ найдутся два базиса такие, что T является матрицей перехода от первого базиса ко второму.*

Доказательство.

1° Матрица перехода $T_{e\to e'}$ состоит из координатных столбцов $[e'_1]_e$, $[e'_2]_e$, $\ldots$, $[e'_n]_e$ векторов базиса. Поскольку векторы базиса линейно независимы, то и их координатные столбцы линейно независимы. Поэтому определитель матрицы перехода не равен нулю и, следовательно, матрица перехода невырождена.

2° Умножим обе части равенства $e' = eT_{e\to e'}$ на обратную матрицу $T^{-1}_{e\to e'}$. Получим $e'T^{-1}_{e\to e'} = eT_{e\to e'}T^{-1}_{e\to e'}$. Произведение $T_{e\to e'}T^{-1}_{e\to e'}$ равно единичной матрице I_n, тогда $e'T^{-1}_{e\to e'} = eI_n$ или $e'T^{-1}_{e\to e'} = e$. Сравнив это равенство с (73), получаем, что $T^{-1}_{e\to e'}$ — матрица перехода от базиса e' к базису e.

3° Возьмём произвольный базис $e = (e_1, e_2, \ldots, e_n)$ и по формулам (71) определим векторы $e'_1, e'_2, \ldots, e'_n$. Координатные столбцы $[e'_j]_e$, $j = \overline{1,n}$ в базисе e совпадают со столбцами матрицы T. Невырожденность матрицы T эквивалентна линейной независимости столбцов, а она эквивалентна линейной независимости векторов $e'_1, e'_2, \ldots, e'_n$. Тогда $e' = (e'_1, e'_2, \ldots, e'_n)$ — базис, и из формул (71) следует, что T является матрицей перехода от базиса e к базису e'. ■

Найдём связь координат произвольного вектора в двух базисах.

Теорема 4.6.2. (Связь координат вектора в двух базисах)

Координаты вектора x в базисах $e = (e_1, e_2, \ldots, e_n)$ и $e' = (e'_1, e'_2, \ldots, e'_n)$ связаны эквивалентными равенствами

$$[x]_e = T_{e\to e'}\,[x]_{e'}\,, \tag{74}$$

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \ldots & t_{1n} \\ t_{21} & t_{22} & \ldots & t_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ t_{n1} & t_{n2} & \ldots & t_{nn} \end{pmatrix} \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \ldots \\ \alpha'_n \end{pmatrix}, \tag{74'}$$

$$\begin{aligned} \alpha_1 &= t_{11}\alpha'_1 + t_{12}\alpha'_2 + \ldots + t_{1n}\alpha'_n, \\ \alpha_2 &= t_{21}\alpha'_1 + t_{22}\alpha'_2 + \ldots + t_{2n}\alpha'_n, \\ &\ldots \\ \alpha_n &= t_{n1}\alpha'_1 + t_{n2}\alpha'_2 + \ldots + t_{nn}\alpha'_n, \end{aligned} \tag{74''}$$

где $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$, $[x]_{e'} = \begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \ldots \\ \alpha'_n \end{pmatrix}$ *— координатные столбцы вектора x в базисах e и e' соответственно, $T_{e\to e'}$ — матрица перехода* (72) *от базиса e к базису e'.*

Доказательство. Из формулы (68) следуют два разложения $x = e[x]_e$ и $x = e'[x]_{e'}$ вектора x по базису e и e' соответственно. Из равенств $x = e'[x]_{e'}$ и $e' = eT_{e\to e'}$ получим $x = e'[x]_{e'} = (eT_{e\to e'})[x]_{e'} = e(T_{e\to e'}[x]_{e'})$. Итак, получено два разложения вектора x по базису e: $x = e[x]_e$ и $x = e(T_{e\to e'}[x]_{e'})$. В силу единственности координат выполнено равенство координатных столбцов $[x]_e = T_{e\to e'}[x]_{e'}$. Отсюда получаются подробные координатные равенства $(74')$ и $(74'')$. ■

Равенства (74) – $(74'')$ называют *формулы преобразования координат.*

Задача 4.6.1. *Как изменится матрица перехода от первого базиса ко второму, если* 1) *в первом базисе поменять местами два вектора,* 2) *во втором базисе поменять местами два вектора,* 3) *вместо m-го вектора первого базиса взять сумму m-го вектора первого базиса и утроенного k-го вектора первого базиса,* 4) *вместо m-го вектора*

второго базиса взять разность m-го вектора второго базиса и утроенного k-го вектора второго базиса, 5) *заменить первый базис на второй, а второй — на первый.*

Задача 4.6.2. *Найдите все* 2×2*-матрицы перехода от первого базиса ко второму, которые совпадают с матрицей перехода от второго базиса к первому.*

Задача 4.6.3. *В вещественном линейном пространстве в некотором базисе заданы координаты двух систем векторов* e_1, e_2, e_3; e'_1, e'_2, e'_3 *соответственно:* $(-1,\ 3,\ 3)$, $(-3,\ 1,\ 2)$, $(4,\ 3,\ 1)$; $(1,\ -2,\ -4)$, $(1,\ 2,\ -1)$, $(-1,\ 1,\ 3)$. *Докажите, что каждая система является базисом. Найдите матрицу перехода от первого базиса ко второму и связь между координатами одного и того же вектора в этих двух базисах.*

Р е ш е н и е.

I. Каждый заданный вектор имеет 3 координаты, поэтому заданы векторы в трёхмерном линейном пространстве. Любые три линейно независимых вектора в этом пространстве линейно независимы.

Для исследования линейной независимости векторов e_1, e_2, e_3 исследуем линейную независимость их координатных столбцов, из которых составим матрицу $A = \begin{pmatrix} -1 & -3 & 4 \\ 3 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix}$. Для вычисления ранга упростим матрицу элементарными преобразованиями строк.

$$A = \begin{pmatrix} -1 & -3 & 4 \\ 3 & 1 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{matrix} \\ 3s^1 + s^2 \\ 3s^1 + s^3 \end{matrix} \to \begin{pmatrix} -1 & -3 & 4 \\ 0 & -8 & 15 \\ 0 & -7 & 13 \end{pmatrix} \begin{matrix} -s^1 \\ -s^3 + s^2 \\ \\ \end{matrix} \to$$

$$\to \begin{pmatrix} 1 & 3 & -4 \\ 0 & -1 & 2 \\ 0 & -7 & 13 \end{pmatrix} \begin{matrix} \\ -s^2 \\ -7s^2 + s^3 \end{matrix} \to \begin{pmatrix} 1 & 3 & -4 \\ 0 & 1 & -2 \\ 0 & 0 & -1 \end{pmatrix} \begin{matrix} -4s^3 + s^1 \\ -2s^3 + s^2 \\ -s^3 \end{matrix} \to$$

$$\to \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} -3s^2 + s^1 \\ \\ \\ \end{matrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ранг исходной и всех преобразованных матриц равен 3, т.к. ранг не меняется при элементарных преобразованиях и определитель последней матрицы $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$ является ненулевым минором порядка 3 и нет ненулевых миноров большего порядка. Определитель матрицы A является базисным минором и по теореме о базисном миноре базисные столбцы матрицы линейно независимы. Следовательно, векторы e_1, e_2, e_3 линейно независимы и образуют базис.

б) Для исследования линейной независимости векторов e'_1, e'_2, e'_3 исследуем линейную независимость их координатных столбцов, из которых составим матрицу $B = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 2 & 1 \\ -4 & -1 & 3 \end{pmatrix}$. Для вычисления ранга упростим матрицу B элементарными преобразованиями строк.

$$B = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 2 & 1 \\ -4 & -1 & 3 \end{pmatrix} \begin{matrix} \\ 2s^1 + s^2 \\ 4s^1 + s^3 \end{matrix} \to \begin{pmatrix} 1 & 1 & -1 \\ 0 & 4 & -1 \\ 0 & 3 & -1 \end{pmatrix} -s^3 + s^2 \to$$

$$\to \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 3 & -1 \end{pmatrix} \begin{matrix} -s^2 + s^1 \\ \\ -3s^2 + s^3 \end{matrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Ранг исходной и всех преобразованных матриц равен 3, т.к. ранг не меняется при элементарных преобразованиях и определитель последней матрицы $\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -1 \neq 0$ является ненулевым минором порядка 3 и нет ненулевых миноров большего порядка. Определитель матрицы B является базисным минором и по теореме о базисном миноре базисные столбцы матрицы линейно независимы. Следовательно, векторы e'_1, e'_2, e'_3 линейно независимы и образуют базис.

II 1) Для нахождения разложений

$$e'_1 = t_{11}e_1 + t_{21}e_2 + t_{31}e_3 \text{ или } t_{11}\begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix} + t_{21}\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} + t_{31}\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ -2 \\ -4 \end{pmatrix}$$

с неизвестными t_{11}, t_{21}, t_{31},

$$e_2' = t_{12}e_1 + t_{22}e_2 + t_{32}e_3 \text{ или } t_{12}\begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix} + t_{22}\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} + t_{32}\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

с неизвестными t_{12}, t_{22}, t_{32},

$$e_2' = t_{13}e_1 + t_{23}e_2 + t_{33}e_3 \text{ или } t_{13}\begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix} + t_{23}\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} + t_{33}\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} =$$

$$\begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$$

с неизвестными t_{13}, t_{23}, t_{33},

можно решить три системы с матрицей A и столбцами свободных членов — координатными столбцами для векторов e_1', e_2', e_3' соответственно или одно матричное уравнение $AT = B$.

Решаем матричное уравнение $AT = B$. Составим и преобразуем расширенную матрицу $(A|B)$ такими же преобразованиями, которые делали в Ia)

$$\left(\begin{array}{ccc|ccc} -1 & -3 & 4 & 1 & 1 & -1 \\ 3 & 1 & 3 & -2 & 2 & 1 \\ 3 & 2 & 1 & -4 & -1 & 3 \end{array}\right) \begin{array}{l} \\ 3s^1 + s^2 \\ 3s^1 + s^3 \end{array} \to \left(\begin{array}{ccc|ccc} -1 & -3 & 4 & 1 & 1 & -1 \\ 0 & -8 & 15 & 1 & 5 & -2 \\ 0 & -7 & 13 & -1 & 2 & 0 \end{array}\right) \begin{array}{l} -s^1 \\ -s^3 + s^2 \\ \\ \end{array} \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 3 & -4 & -1 & -1 & 1 \\ 0 & -1 & 2 & 2 & 3 & -2 \\ 0 & -7 & 13 & -1 & 2 & 0 \end{array}\right) \begin{array}{l} \\ -s^2 \\ -7s^2 + s^3 \end{array} \to \left(\begin{array}{ccc|ccc} 1 & 3 & -4 & -1 & -1 & 1 \\ 0 & 1 & -2 & -2 & -3 & 2 \\ 0 & 0 & -1 & -15 & -19 & 14 \end{array}\right) \begin{array}{l} -4s^3 + s^1 \\ -2s^3 + s^2 \\ -s^3 \end{array} \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & 59 & 75 & -55 \\ 0 & 1 & 0 & 28 & 35 & -26 \\ 0 & 0 & 1 & 15 & 19 & -14 \end{array}\right) \begin{array}{l} -3s^2 + s^1 \\ \\ \\ \end{array} \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -25 & -30 & 23 \\ 0 & 1 & 0 & 28 & 35 & -26 \\ 0 & 0 & 1 & 15 & 19 & -14 \end{array}\right)$$

2) Матрица перехода $T = \begin{pmatrix} -25 & -30 & 23 \\ 28 & 35 & -26 \\ 15 & 19 & -14 \end{pmatrix}$.

III. Связь между координатами произвольного вектора x в этих

двух базисах задаётся формулами преобразования координат:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -25 & -30 & 23 \\ 28 & 35 & -26 \\ 15 & 19 & -14 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$$

или

$$\begin{aligned} x_1 &= -25x_1' - 30x_2' + 23x_3', \\ x_2 &= 28x_1' + 35x_2' - 26x_3', \\ x_3 &= 15x_1' + 19x_2' - 14x_3' \end{aligned}$$

где вектор x имеет в первом базисе координаты $(x_1,\, x_2,\, x_3)$ и во втором базисе — $(x_1',\, x_2',\, x_3')$.

4.7. Линейная оболочка векторов

В этом разделе рассматривается один из способов построения линейные подпространства линейного пространства E над полем P.

Определение 1. **Линейной оболочкой** *векторов*

$$e_1,\, e_2,\, \ldots,\, e_k \in E \tag{75}$$

называется множество всевозможных линейных комбинаций этих векторов. Обозначение: $\mathrm{Lin}(e_1,\, e_2,\, \ldots,\, e_k)$.

Итак, $\mathrm{Lin}(e_1,\, e_2,\, \ldots,\, e_k) = \{x\colon x = \sum\limits_{i=1}^{k} \alpha_i e_i,\, \alpha_i \in P,\, i = \overline{1,k}\}$.

Теорема 4.7.1.

Если $e_1,\, e_2,\, \ldots,\, e_k$ *— система векторов пространства* E, *то их линейная оболочка является конечномерным подпространством в* E. *Базисом ненулевой линейной оболочки* $\mathrm{Lin}(e_1,\, e_2,\, \ldots,\, e_k)$ *является любая база системы векторов* $e_1,\, e_2,\, \ldots,\, e_k$, *а размерность линейной оболочки равна рангу системы векторов.*

Доказательство.

Обозначим $L = \mathrm{Lin}(e_1, e_2, \ldots, e_k)$. L — это непустое подмножество пространства, т. к. $\theta = \sum\limits_{i=1}^{k} 0e_i \in L$. Применим к L критерий подпространства.

а) Если $f, g \in L$, то $f = \sum\limits_{i=1}^{k} \alpha_i e_i$, $g = \sum\limits_{i=1}^{k} \beta_i e_i$, где $\alpha_i, \beta_i \in P$, $i = \overline{1,k}$. Тогда $f + g = \sum\limits_{i=1}^{k} (\alpha_i + \beta_i) e_i$, где $\alpha_i + \beta_i \in P$, $i = \overline{1,k}$, то есть $f + g \in L$.

б) Если $f \in L$ и $\gamma \in P$, то $f = \sum\limits_{i=1}^{k} \alpha_i e_i$, где $\alpha_i \in P, i = \overline{1,k}$. Тогда $\gamma f = \sum\limits_{i=1}^{k} (\gamma \alpha_i) e_i$, где $\gamma \alpha_i \in P$, $i = \overline{1,k}$, то есть $\gamma f \in L$.

Из а) и б) следует, что линейная оболочка $L = \mathrm{Lin}(e_1, e_2, \ldots, e_k)$ является подпространством в E.

Если все векторы e_i нулевые, то $L = \theta$ и $\dim(L) = 0$.

Пусть теперь не все векторы e_i нулевые. В этом случае система векторов $e_1, e_2, \ldots, e_k$ содержит базу $f_1, \ldots, f_r$. Любой вектор x линейно выражается через векторы $e_1, e_2, \ldots, e_k$, а они линейно выражаются через векторы базы $f_1, \ldots, f_r$. Поэтому любой вектор x линейно выражается через векторы базы, причём векторы базы линейно независимы. Тогда векторы базы $f_1, \ldots, f_r$ образуют базис подпространства L и, следовательно, $\dim(L) = r = rg(e_1, e_2, \ldots, e_k)$. ■

Теорема 4.7.2. *Пусть L_1, L_2 — линейные подпространства линейного пространства E.*

1° *Если $L_1 \subseteq L_2$, то $\dim(L_1) \leqslant \dim(L_2)$ (монотонность размерностей).*

2° *Если $L_1 \subseteq L_2$, $\dim(L_1) = \dim(L_2) < +\infty$, то $L_1 = L_2$.*

Доказательство. 1° Любая линейно независимая система векторов из L_1 в силу $L_1 \subseteq L_2$ является линейно независимой системой в L_2. Поэтому неравенство $\dim(L_1) > \dim(L_2)$ невозможно и, следовательно, $\dim(L_1) \leqslant \dim(L_2)$.

2° При $\dim(L_1) = \dim(L_2) = 0$ оба подпространства нулевые и утверждение $L_1 = L_2$ верно.

При $\dim(L_1) = \dim(L_2) = n > 0$ выберем n линейно независимых векторов $e_1, e_2, \ldots, e_n$ из L_1. Поскольку $\dim(L_1) = n$ эти векторы образуют базис в L_1. В силу $L_1 \subseteq L_2$ и $\dim(L_2) = n$ эти n линейно независимых векторов принадлежат L_2 и, следовательно, они об-

разуют базис в L_2. Используя полученные факты, докажем включение $L_2 \subseteq L_1$. Возьмём любой вектор x из L_2 и разложим его по базису: $x = \sum\limits_{i=1}^{n} \alpha_i e_i$, $\alpha_i \in P$, $i = \overline{1,n}$. Все векторы e_i, $i = \overline{1,n}$ принадлежат L_1, и тогда их линейная комбинация $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in L_1$. Итак, доказано включение $L_2 \subseteq L_1$, что вместе с условием $L_1 \subseteq L_2$ даёт равенство $L_1 = L_2$. ■

Следствие. *Линейная оболочка базиса пространства совпадает с этим пространством.*

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ базис пространства E и тогда $\dim(E) = n$. Из утверждений теоремы 4.7.1 следует, что линейная оболочка $L = \mathrm{Lin}(e_1, e_2, \ldots, e_n)$ является подпространством в E и базисом L является базис пространства, поэтому $\dim(L) = n$. Итак, $L \subseteq E$, $\dim(L) = \dim(E)$ и тогда из утверждения 2° теоремы 4.7.2 получаем $L = E$. ■

Задача 4.7.1. *Докажите, что если для множеств векторов выполнено включение* $\{e_1, e_2, \ldots, e_k\} \subseteq \{g_1, g_2, \ldots, g_m\}$, *то для линейных оболочек этих множеств выполнено включение* $\mathrm{Lin}(e_1, e_2, \ldots, e_k) \subseteq \mathrm{Lin}(g_1, g_2, \ldots$

4.8. Сумма и пересечение подпространств

В линейном пространстве E над полем P рассмотрим операции пересечения множеств и сложения множеств[4]. С помощь этих операций можно по заданным подпространствам построить другие подпространства.

Определение 1. 1) *Пусть L_δ — подпространство в E для каждого индекса $\delta \in \Phi$, где Φ — множество индексов.*[5]

Пересечением подпространств L_δ, $\delta \in \Phi$ называется множество, состоящее из тех векторов, которые принадлежат всем подпространствам L_δ.

Обозначение: $\bigcap\limits_{\delta \in \Phi} L_\delta$.

[4] Не путать её с операцией объединения множеств.
[5] Φ — конечное или бесконечное множество.

2) *Пусть* $L_1, L_2, \ldots, L_k$ *— подпространства в* E. *Суммой подпространств* L_i, $i = \overline{1,k}$ *называется множество тех векторов* $x \in E$, *которые представимы в виде*

$$x = x_1 + x_2 + \ldots + x_k, \tag{76}$$

где $x_i \in L_i$, $i = \overline{1,k}$.

Обозначение: $L_1 + L_2 + \ldots + L_k$ *или* $\sum\limits_{i=1}^{k} L_i$.

Итак, $\sum\limits_{i=1}^{k} L_k = \{x \,|\, x = x_1 + x_2 + \ldots + x_k,\ x_i \in L_i,\ i = \overline{1,k}\}$.

Равенство (76) *называется разложением вектора* x *по подпространствам* $L_1, L_2, \ldots, L_k$.

Очевидно, что если среди данных подпространств есть нулевые, то пересечением таких подпространств будет нулевое подпространство, сумма данных подпространств не изменится при удалении всех нулевых подпространств.

Теорема 4.8.1.

1° *Пересечение любого множества подпространств является подпространством в* E.

2° *Сумма подпространств является подпространством в* E. *Базисом суммы ненулевых конечномерных подпространств является любая база системы векторов, полученной объединением базисов этих подпространств.*

Доказательство. Используем критерий подпространства.

1° Пусть L_δ подпространство в E при каждом $\delta \in \Phi$ и $L = \bigcap\limits_{\delta \in \Phi} L_\delta$. Заметим, что пересечение подпространств непусто, т.к. нулевой вектор принадлежит каждому подпространству и, следовательно, их пересечению.

а) Для $x, y \in L$ проверим, что $x + y \in L$.

Если $x, y \in L$, то при всех $\delta \in \Phi$ выполнены условия $x, y \in L_\delta$. При каждом $\delta \in \Phi$ L_δ — подпространство, поэтому $x + y \in L_\delta$ при каждом $\delta \in \Phi$. Следовательно, $x + y$ принадлежит пересечению всех подпространств L_δ, т. е. $x + y \in L$.

б) Если $x \in L$, $\alpha \in P$, то при всех $\delta \in \Phi$ выполнены условия $x \in L_\delta$. При каждом $\delta \in \Phi$ L_δ — подпространство, поэтому вектор $\alpha x \in L_\delta$ при каждом $\delta \in \Phi$. Следовательно, αx принадлежит пересечению всех подпространств L_δ, т.е. $\alpha x \in L$.

Из а) и б) следует, что $L = \bigcap\limits_{\delta \in \Phi} L_\delta$ — подпространство.

2° Пусть $L_1, L_2, \ldots, L_k$ подпространства в E и $L = \sum\limits_{i=1}^{k} L_i$.

а) Для $x, y \in L$ проверим, что $x + y \in L$.

Если $x, y \in L$, то выполнены условия $x = \sum\limits_{i=1}^{k} x_i$, $y = \sum\limits_{i=1}^{k} y_i$, где $x_i \in L_i$, $y_i \in L_i$, $i = \overline{1,k}$. Тогда для $x + y$ получаем разложение $x + y = \sum\limits_{i=1}^{k} x_i + \sum\limits_{i=1}^{k} y_i = \sum\limits_{i=1}^{k} (x_i + y_i)$. Здесь $x_i + y_i \in L_i$, т. к. L_i — подпространство и $x_i, y_i \in L_i$. Следовательно, $x + y \in L$.

б) Для $x \in L$, $\alpha \in P$ проверим, что $\alpha x \in L$.

Если $x \in L$, то выполнено условие $x = \sum\limits_{i=1}^{k} x_i$, где $x_i \in L_i$, $i = \overline{1,k}$. Тогда для αx получаем разложение $\alpha x = \alpha \sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} \alpha x_i$. Здесь $\alpha x_i \in L_i$, т. к. L_i — подпространство и $x_i \in L_i$. Следовательно, $\alpha x \in L$.

Из а) и б) следует, что сумма подпространств $L = \sum\limits_{i=1}^{k} L_i$ — подпространство.

Пусть $L = \sum\limits_{i=1}^{k} L_i$, где L_i — ненулевые конечномерные подпространства. Обозначим через F_i, $i = \overline{1,k}$ базис подпространства L_i и через F объединение базисов $\bigcup\limits_{i=1}^{k} F_i$.

Покажем, что подпространство L есть линейная оболочка векторов F, т.е. $L = \mathrm{Lin}(F)$. Для этого докажем включения $L \subseteq \mathrm{Lin}(F)$ и $\mathrm{Lin}(F) \subseteq L$.

Проверим $L \subseteq \mathrm{Lin}(F)$. Пусть $x \in L = \sum\limits_{i=1}^{k} L_i$, тогда $x = \sum\limits_{i=1}^{k} x_i$, где $x_i \in L_i$, $i = \overline{1,k}$. Вектор x_i есть линейная комбинация векторов базиса F_i, тогда при сложении получим, что $x = \sum\limits_{i=1}^{k} x_i$ есть линейная комбинация векторов из объединения всех базисов F, т.е $x \in \mathrm{Lin}(F)$.

Проверим $\mathrm{Lin}(F) \subseteq L$. Пусть $x \in \mathrm{Lin}(F)$, т. е. x равен линейной комбинации векторов из $F = \bigcup\limits_{i=1}^{k} F_i$. Тогда в этой линейной комбинации сгруппируем 1) слагаемые с векторами из F_1 и обозначим их сумму x_1, 2) слагаемые с векторами из F_2 и обозначим их сумму x_2, и т. д. ..., k) слагаемые с векторами из F_k и обозначим их сумму x_k. Получим, разложение $x = \sum\limits_{i=1}^{k} x_i$, где $x_i \in L_i$, $i = \overline{1,k}$, так как x_i разложено по базису F_i подпространства L_i. Следовательно, $x = \sum\limits_{i=1}^{k} x_i \in L = \sum\limits_{i=1}^{k} L_i$, т. е. $\mathrm{Lin}(F) \subseteq L$.

Из включений $L \subseteq \mathrm{Lin}(F)$ и $\mathrm{Lin}(F) \subseteq L$ следует равенство $L = \mathrm{Lin}(F)$. Тогда базисом суммы подпространств $L = \sum\limits_{i=1}^{k} L_i$ является базис линейной оболочки $\mathrm{Lin}(F)$. А по теореме 4.7.1 базисом $\mathrm{Lin}(F)$ является любая база системы векторов F. ■

Следствие. *Пусть L_1 и L_2 ненулевые конечномерные подпространства линейного пространства E и $L_1 \cap L_2 = \theta$. Тогда объединение векторов базисов этих подпространств является базисом суммы $L_1 + L_2$.*

Доказательство. Пусть $e_1, \ldots, e_m$ и $f_1, \ldots, f_k$ — базисы в L_1 и L_2 соответственно.

Покажем, что объединение базисов

$$e_1, \ldots, e_m, f_1, \ldots, f_k \tag{77}$$

является базисом $L_1 + L_2$.

Проверим линейную независимость векторов (77). Пусть линейная комбинация векторов (77) равна θ : $\sum\limits_{i=1}^{m} \alpha_i e_i + \sum\limits_{i=1}^{k} \beta_i f_i = \theta$. Тогда $\sum\limits_{i=1}^{m} \alpha_i e_i = \sum\limits_{i=1}^{k} (-\beta_i) f_i$. Обозначив $h = \sum\limits_{i=1}^{m} \alpha_i e_i = \sum\limits_{i=1}^{k} (-\beta_i) f_i$, имеем h разложен по базису в L_1 и h разложен по базису в L_2, т.е. $h \in L_1 \cap L_2$. Из условия $L_1 \cap L_2 = \theta$ следует, что $h = \theta$. Поэтому $\sum\limits_{i=1}^{m} \alpha_i e_i = \theta$, $\sum\limits_{i=1}^{k} (-\beta_i) f_i = \theta$ и из линейной независимости векторов каждого базиса получаем, что все коэффициенты $\alpha_i = 0$ и $\beta_i = 0$. Поэтому линейная комбинация $\sum\limits_{i=1}^{m} \alpha_i e_i + \sum\limits_{i=1}^{k} \beta_i f_i$ — тривиальная. Тогда векторы (77) линейно независимы

и по теореме 4.8.1 образуют базис $L_1 + L_2$. ■

Задача 4.8.1. *а) Докажите, что* $L_1 \subseteq L_1 + L_2$ *для любых подпространств* L_1 *и* L_2.

б) Пусть L_1, L_2, M_1, M_2 *подпространства линейного пространства* E *и* $L_1 \subseteq L_2$, $M_1 \subseteq M_2$. *Как связаны подпространства* $L_1 + M_1$ *и* $L_2 + M_2$.

в) Докажите, что
$\mathrm{Lin}(e_1, \ldots, e_k) + \mathrm{Lin}(g_1, \ldots, g_m) = \mathrm{Lin}(e_1, \ldots, e_k, g_1, \ldots, g_m)$.

4.9. Размерность суммы двух подпространств

В этом разделе продолжим изучать подпространства линейного пространства E над полем P. Здесь формулируется и доказывается теорема о размерности суммы двух подпространств.

Теорема 4.9.1.

Если L_1 *и* L_2 *любые конечномерные подпространства линейного пространства* E, *то размерность суммы подпространств равна сумме размерностей этих подпространств минус размерность их пересечения, т.е.*

$$\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2) - \dim(L_1 \cap L_2). \qquad (78)$$

Доказательство. Из определений пересечения и суммы подпространств следует, что $L_1 \cap L_2 \subseteq L_1 \subseteq L_1 + L_2$, $L_1 \cap L_2 \subseteq L_2 \subseteq L_1 + L_2$. Тогда из теорем 4.7.2 и 4.8.1 получаем, что все указанные множества являются конечномерными подпространствами линейного пространства E.

Если $L_1 \cap L_2 = \theta$, то $\dim(L_1 \cap L_2) = 0$. При этом из следствия теоремы 4.8.1 следует, что $\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2)$, т.е. равенство (78) выполнено.

Рассмотрим случай $L_1 \cap L_2 \neq \theta$.

При $L_1 = L_1 \cap L_2$ имеем $L_1 \subseteq L_2$. Тогда $L_1 + L_2 \subseteq L_2 + L_2 = L_2$. Поскольку верно включение $L_2 \subseteq L_1 + L_2$, то получаем $L_2 = L_1 + L_2$. Тогда формула (78) верна, т. к. имеет вид $\dim(L_2) = \dim(L_1 \cap L_2) +$

$\dim(L_2) - \dim(L_1 \cap L_2)$. Аналогичные рассуждения проводятся при $L_2 = L_1 \cap L_2$.

Остаётся рассмотреть наиболее сложный случай $L_1 \cap L_2 \neq \theta$, $L_1 \neq \neq L_1 \cap L_2$, $L_2 \neq L_1 \cap L_2$. Результат получим с помощью конструирования базисов в $L_1 \cap L_2$, L_1, L_2, $L_1 + L_2$.

Возьмём базис $t = (e_1, \ldots, e_m)$ в $L_1 \cap L_2$. Эти векторы линейно независимы и в силу $L_1 \cap L_2 \subseteq L_1$ принадлежат L_1. Так как $L_1 \cap L_2 \neq L_1$, векторы из t не образуют базис подпространства L_1, но их можно дополнить до базиса $u = (e_1, \ldots, e_m, f_1 \ldots, f_k)$ в L_1.

Так как $t \subseteq L_1 \cap L_2 \subseteq L_2$, $L_1 \cap L_2 \neq L_2$, аналогично дополним векторы из t до базиса $v = (e_1, \ldots, e_m, g_1 \ldots, g_n)$ в L_2.

Объединив базисы u и v получим совокупность векторов

$$w = (e_1, \ldots, e_m, f_1, \ldots, f_k, g_1, \ldots, g_n). \tag{79}$$

Покажем, что векторы (79) линейно независимы.

Пусть линейная комбинация векторов (79) равна θ :

$$\sum_{i=1}^{m} \alpha_i e_i + \sum_{i=1}^{k} \beta_i f_i + \sum_{i=1}^{n} \gamma_i g_i = \theta. \tag{80}$$

Отсюда $\sum\limits_{i=1}^{k} \beta_i f_i = \sum\limits_{i=1}^{m} (-\alpha_i) e_i + \sum\limits_{i=1}^{n} (-\gamma_i) g_i$. Обозначив $h = \sum\limits_{i=1}^{k} \beta_i f_i = = \sum\limits_{i=1}^{m} (-\alpha_i) e_i + \sum\limits_{i=1}^{n} (-\gamma_i) g_i$, имеем h разложен по базису u в L_1 и h разложен по базису v в L_2, т.е. $h \in L_1 \cap L_2$. Отсюда вектор h имеет разложение по базису t в $L_1 \cap L_2$: $h = \sum\limits_{i=1}^{m} \delta_i e_i$. Тогда для вектора h имеем два разложения по базису v в L_2 : $h = \sum\limits_{i=1}^{m} \delta_i e_i + \sum\limits_{i=1}^{n} 0 g_i$ и $h = \sum\limits_{i=1}^{m} (-\alpha_i) e_i + \sum\limits_{i=1}^{n} (-\gamma_i) g_i$. В силу единственности разложения по базису имеем равенства всех соответствующих коэффициентов $\delta_i = -\alpha_i$ и $0 = -\gamma_i$. Отсюда все коэффициенты $\gamma_i = 0$, и поэтому равенство (80) принимает вид $\sum\limits_{i=1}^{m} \alpha_i e_i + \sum\limits_{i=1}^{k} \beta_i f_i = \theta$. Здесь в силу линейной независимости векторов базиса $u = (e_1, \ldots, e_m, f_1, \ldots, f_k)$ все коэффициенты $\alpha_i = 0$ и $\beta_i = 0$. Таким образом, равенству (80) удовлетворяет только тривиальная линейная комбинация, т. е. векторы (79) линейно независи-

мы. При этом они являются базой системы векторов (79) и по теореме 4.8.1 образуют базис $L_1 + L_2$.

Размерность подпространств равна количеству векторов базиса:

$$\dim(L_1 + L_2) = m + k + n, \qquad \dim(L_1) = m + k,$$
$$dim(L_2) = m + n, \qquad \dim(L_1 \cap L_2) = m.$$

Равенство (78) для этих величин выполнено: $m + k + n = (m + k) + (m + n) - m$. ■

Задача 4.9.1. *а) Докажите, что* $\dim(L_1+L_2) \leqslant \dim(L_1)+\dim(L_2)$ *для любых подпространств* L_1 *и* L_2.

б) Докажите, что $\dim(L_1+L_2+\ldots+L_k) \leqslant \dim(L_1)+\dim(L_2)+\ldots+\dim(L_k)$ *для любых подпространств* $L_1, L_2, \ldots, L_k$, $k > 2$.

Задача 4.9.2. *В вещественном линейном пространстве в некотором базисе заданы координаты 2-х систем векторов:*
$a_1(5, 8, -10, -2)$, $a_2(-1, -1, 2, 7)$, $a_3(2, 2, -4, -8)$, $a_4(-1, -1, 2, 4)$ *и* $b_1(1, 7, -10 - 5)$, $b_2(2, 5, -8, -4)$, $b_3(-3, 2, -2, -1)$, $b_4(4, 1, 0, 0)$.
Линейные оболочки этих систем векторов образуют два подпространства. Найдите базисы и размерности каждого подпространства, а также их суммы и пересечения.

Решение задач.

4.9.2 Обозначим линейные оболочки векторов $L_1 = Lin(a_1, a_2, a_3, a_4)$ и $L_2 = Lin(b_1, b_2, b_3, b_4)$.

1) Ищем базис и размерность L_1.

Для исследования линейной зависимости векторов a_1, a_2, a_3, a_4 исследуем линейную зависимость их координатных столбцов, из которых составим матрицу: $A = \begin{pmatrix} 5 & -1 & 2 & -1 \\ 8 & -1 & 2 & -1 \\ -10 & 2 & -4 & 2 \\ -2 & 7 & -8 & 4 \end{pmatrix}$. Упростим матрицу A элементарными преобразованиями строк и переменой порядка столбцов:

$$\begin{pmatrix} 5 & -1 & 2 & -1 \\ 8 & -1 & 2 & -1 \\ -10 & 2 & -4 & 2 \\ -2 & 7 & -8 & 4 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & -1 & 2 & 5 \\ -1 & -1 & 2 & 8 \\ 2 & 2 & -4 & -10 \\ 4 & 7 & -8 & -2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & -1 & 2 & 5 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 18 \end{pmatrix} \Leftrightarrow$$

меняем 1 и 4 столбцы, $r_2 := r_2 - r_1$, $r_3 := r_3 + 2r_1$, $r_4 := r_4 + 4r_1$, $r_4 := r_4 - 6r_2$,

$$\Leftrightarrow \begin{pmatrix} -1 & -1 & 2 & 5 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & -1 & 2 & 5 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & -1 & 5 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$(r_2, r_3, r_4) := (r_4, r_2, r_3)$ меняем 3 и 4 столбцы

В последней матрице минор $\begin{vmatrix} -1 & -1 & 5 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{vmatrix} = -9$ является ненулевым минором порядка 3, а ненулевых миноров большего порядка нет. Значит ранг этой матрицы равен трём, а указанный минор является базисным минором. Ему соответствует в матрице A базисный минор, расположенный в строках с номерами 1, 2, 3 и столбцах с номерами 4, 2, 1 (проверьте!). По теореме о базисном миноре базисные столбцы матрицы линейно независимы и через них линейно выражаются все столбцы. Следовательно, векторы a_4, a_2, a_1 линейно независимы и через них линейно выражаются векторы a_1, a_2, a_3, a_4, а значит и все векторы в линейной оболочке L_1. Поэтому векторы a_4, a_2, a_1 (можно изменить их порядок) образуют базис в L_1, и тогда размерность $\dim(L_1) = 3$.

2) Ищем базис и размерность L_2.

Из координатных столбцов векторов b_1, b_2, b_3, b_4 составим матрицу:

$$B = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 7 & 5 & 2 & 1 \\ -10 & -8 & -2 & 0 \\ -5 & -4 & -1 & 0 \end{pmatrix}.$$

Упростим матрицу элементарными преобразованиями строк и переменой столбцов. Сначала меняем 1-й и 4-й столбцы и 1-ю и 2-ю строки. $\Leftrightarrow$

$$\Leftrightarrow \begin{pmatrix} 1 & 5 & 2 & 7 \\ 4 & 2 & -3 & 1 \\ 0 & -8 & -2 & -10 \\ 0 & -4 & -1 & -5 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 5 & 2 & 7 \\ 0 & -18 & -11 & -27 \\ 0 & 4 & 1 & 5 \\ 0 & -4 & -1 & -5 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 5 & 2 & 7 \\ 0 & -18 & -11 & -27 \\ 0 & 4 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$r_2 := r_2 - 4r_1$, $r_3 := -r_3/2$ $\quad$ $r_4 := r_4 + r_3$

В последней матрице минор $\begin{vmatrix} 1 & 5 & 2 \\ 0 & -18 & -11 \\ 0 & 4 & 1 \end{vmatrix} = \begin{vmatrix} -18 & -11 \\ 4 & 1 \end{vmatrix} = -18+44 =$ $= 26$ является ненулевым минором порядка 3 и нет ненулевых миноров большего порядка. Значит ранг этой матрицы равен 3 и указанный минор является базисным. Ему соответствует в матрице B базисный минор, расположенный в в строках с номерами $1,\ 2,\ 3$ и столбцах с номерами $4,\ 2,\ 1$ (проверьте!). По теореме о базисном миноре базисные столбцы матрицы линейно независимы и через них линейно выражаются все столбцы. Следовательно, векторы $b_4,\ b_2,\ b_3$ линейно независимы и через них линейно выражаются векторы $b_1, b_2, b_3, b_4,$ а значит и все векторы в линейной оболочке L_2. Поэтому векторы $b_4,\ b_2,\ b_3$ (можно изменить их порядок) образуют базис в $L_2,$ и тогда размерность $\dim(L_2) = 3.$

3) Ищем базис и размерность $L_1 + L_2.$

По теореме 4.8.1 $L_1 + L_2 = Lin(a_4, a_2, a_1, b_4, b_2, b_3),$ причём базисом $L_1 + L_2$ является база системы векторов $a_4, a_2, a_1, b_4, b_2, b_3.$ Исследуем линейную зависимость их координатных столбцов, из которых составим матрицу

$$C = \begin{pmatrix} -1 & -1 & 5 & 4 & 2 & -3 \\ -1 & -1 & 8 & 1 & 5 & 2 \\ 2 & 2 & -10 & 0 & -8 & -2 \\ 4 & 7 & -2 & 0 & -4 & -1 \end{pmatrix}.$$

Упростим матрицу C элементарными преобразованиями строк и переменой столбцов. Сначала меняем $r_2 := r_2 - r_1,\ r_3 := r_3 + 2r_1,\ r_4 := r_4 + 4r_1.$

$$C \Leftrightarrow \begin{pmatrix} -1 & -1 & 5 & 4 & 2 & -3 \\ 0 & 0 & 3 & -3 & 3 & 5 \\ 0 & 0 & 0 & 8 & -4 & -8 \\ 0 & 3 & 18 & 16 & 4 & -13 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & -1 & 5 & 4 & 2 & -3 \\ 0 & 3 & 18 & 16 & 4 & -13 \\ 0 & 0 & 3 & -3 & 3 & 5 \\ 0 & 0 & 0 & 8 & -4 & -8 \end{pmatrix} \Leftrightarrow$$

$(r_2, r_3, r_4) := (r_4, r_2, r_3)$ $\qquad$ $r_1 := -r_1, r_4 := r_4/8$

$$\Leftrightarrow \begin{pmatrix} 1 & 1 & -5 & -4 & -2 & 3 \\ 0 & 3 & 18 & 16 & 4 & -13 \\ 0 & 0 & 3 & -3 & 3 & 5 \\ 0 & 0 & 0 & 1 & -1/2 & -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 1 & -5 & 0 & -4 & -1 \\ 0 & 3 & 18 & 0 & 12 & 3 \\ 0 & 0 & 3 & 0 & 3/2 & 2 \\ 0 & 0 & 0 & 1 & -1/2 & -1 \end{pmatrix} \Leftrightarrow$$

$r_1 := r_1 + 4r_4, r_2 := r_2 - 16r_4,$ $\qquad$ $r_1 := r_1 + 5r_3/3,$

$r_3 := r_3 + 3r_4$ $\qquad$ $r_2 := r_2 - 6r_3$

$$\Leftrightarrow \begin{pmatrix} 1 & 1 & 0 & 0 & -3/2 & 7/3 \\ 0 & 3 & 0 & 0 & 3 & -9 \\ 0 & 0 & 3 & 0 & 3/2 & 2 \\ 0 & 0 & 0 & 1 & -1/2 & -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -5/2 & 16/3 \\ 0 & 1 & 0 & 0 & 1 & -3 \\ 0 & 0 & 1 & 0 & 1/2 & 2/3 \\ 0 & 0 & 0 & 1 & -1/2 & -1 \end{pmatrix}.$$

$r_1 := r_1 - r_2/3$

$r_2 := r_2/3, r_3 := r_3/3$

Минор последней матрицы $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$ является ненулевым минором порядка 4 и нет ненулевых миноров большего порядка. Поэтому ранг последней матрицы равен 4 и этот минор является её базисным минором. Ему соответствует в матрице C базисный минор, расположенный в столбцах с номерами 1, 2, 3, 4, (проверьте!). По теореме о базисном миноре базисные столбцы матрицы линейно независимы и через них линейно выражаются все столбцы. Тогда этим столбцам соответствуют векторы a_4, a_2, a_1, b_4, которые образуют базу системы векторов a_4, a_2, a_1, b_4, b_2, b_3, а векторы базы образуют базис в $L_1 + L_2$. Поэтому $\dim(L_1 + L_2) = 4$.

4) Ищем базис и размерность $L_1 \cap L_2$.

Из формулы $\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2) - \dim(L_1 \cap L_2)$ выразим $\dim(L_1 \cap L_2) = \dim(L_1) + \dim(L_2) - \dim(L_1 + L_2) = 3 + 3 - 4 = 2$. Следовательно, $L_1 \cap L_2$ — двумерное подпространство и нужно найти в нём два линейно независимых вектора, которые дадут базис.

Подпространство $L_1 \cap L_2$ состоит из всех векторов x, которые принадлежат одновременно L_1 и L_2. Поэтому любой вектор $x \in L_1 \cap L_2$ раскладывается по базисам каждого подпространства: $x = \beta_1 a_4 + \beta_2 a_2 + \beta_3 a_1$ и $x = \gamma_1 b_4 + \gamma_2 b_2 + \gamma_3 b_3$, где β_1, β_2, β_3, γ_1, γ_2, γ_3 вещественные числа. Отсюда следует равенство $\beta_1 a_4 + \beta_2 a_2 + \beta_3 a_1 = \gamma_1 b_4 + \gamma_2 b_2 + \gamma_3 b_3$, которое будем рассматривать как уравнение относительно неизвестных β_1, β_2, β_3, γ_1, γ_2, γ_3 для нахождения всех векторов в $L_1 \cap L_2$ и выбора из них базиса. Перенесём все слагаемые в левую часть уравнения

$$\beta_1 a_4 + \beta_2 a_2 + \beta_3 a_1 + (-\gamma_1) b_4 + (-\gamma_2) b_2 + (-\gamma_3) b_3 = \theta,$$

и запишем его в эквивалентном виде для столбцов координат векторов

$$\beta_1 [a_4] + \beta_2 [a_2] + \beta_3 [a_1] + (-\gamma_1)[b_4] + (-\gamma_2)[b_2] + (-\gamma_3)[b_3] = [\theta].$$

Это уравнение эквивалентно системе 6-и уравнений с матрицей C (см. стр. 144), столбцами которой являются координатные столбцы $[a_4]$, $[a_2]$, $[a_1]$, $[b_4]$, $[b_2]$, $[b_3]$. В п.3) проведены элементарные преобразования матрицы с использованием только преобразования строк. Поэтому сразу запишем систему с преобразованной матрицей:

$$\begin{cases} \beta_1 + (-5/2)(-\gamma_2) + (16/3)(-\gamma_3) = 0, \\ \beta_2 + (-\gamma_2) + (-3)(-\gamma_3) = 0, \\ \beta_3 + (1/2)(-\gamma_2) + (2/3)(-\gamma_3) = 0, \\ (-\gamma_1) + (-1/2)(-\gamma_2) + (-1)(-\gamma_3) = 0. \end{cases}$$

Отсюда выражаем базисные неизвестные β_1, β_2, β_3, γ_1, через свободные γ_2, γ_3

и находим общее решение $\begin{cases} \beta_1 = (-5/2)\gamma_2 + (16/3)\gamma_3, \\ \beta_2 = \gamma_2 - 3\gamma_3, \\ \beta_3 = (1/2)\gamma_2 + (2/3)\gamma_3, \\ \gamma_1 = (1/2)\gamma_2 + \gamma_3, \end{cases}$ где γ_2, γ_3 принимают любые значения.

Получим фундаментальную систему решений (ФСР[6]):

1) при $\gamma_2 = 2, \gamma_3 = 0$ находим $\beta_1 = -5$, $\beta_2 = 2$, $\beta_3 = 1$, $\gamma_1 = 1$,
   т.е. $\beta_1 = -5$, $\beta_2 = 2$, $\beta_3 = 1$, $\gamma_1 = 1, \gamma_2 = 2, \gamma_3 = 0$,
2) при $\gamma_2 = 0, \gamma_3 = 3$ находим $\beta_1 = 16$, $\beta_2 = -9$, $\beta_3 = 2$, $\gamma_1 = 3$,
   т.е. $\beta_1 = 16$, $\beta_2 = -9$, $\beta_3 = 2$, $\gamma_1 = 3, \gamma_2 = 0, \gamma_3 = 3$.

Таким образом, с помощью найденных решений мы построим два вектора $x = \beta_1 a_4 + \beta_2 a_2 + \beta_3 a_1 = \gamma_1 b_4 + \gamma_2 b_2 + \gamma_3 b_3$ из пересечения $L_1 \cap L_2$: $x_1 = -5a_4 + 2a_2 + a_1 = b_4 + 2b_2 + 0b_3$ и $x_2 = 16a_4 - 9a_2 + 2a_1 = 3b_4 + 0b_2 + 3b_3$. Для контроля проверим совпадение комбинаций координатных столбцов:

$$-5[a_4]+2[a_2]+[a_1] = -5\begin{pmatrix} -1 \\ -1 \\ 2 \\ 4 \end{pmatrix} + 2\begin{pmatrix} -1 \\ -1 \\ 2 \\ 7 \end{pmatrix} + \begin{pmatrix} 5 \\ 8 \\ -10 \\ -2 \end{pmatrix} = \begin{pmatrix} 5-2+5 \\ 5-2+8 \\ -10+4-10 \\ -20+14-2 \end{pmatrix} = \begin{pmatrix} 8 \\ 11 \\ -16 \\ -8 \end{pmatrix},$$

$$[b_4] + 2[b_2] + 0[b_3] = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 5 \\ -8 \\ -4 \end{pmatrix} + 0\begin{pmatrix} -3 \\ 2 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4+4 \\ 1+10 \\ -16 \\ -8 \end{pmatrix} = \begin{pmatrix} 8 \\ 11 \\ -16 \\ -8 \end{pmatrix},$$

$$16[a_4]-9[a_2]+2[a_1] = 16\begin{pmatrix} -1 \\ -1 \\ 2 \\ 4 \end{pmatrix} - 9\begin{pmatrix} -1 \\ -1 \\ 2 \\ 7 \end{pmatrix} + 2\begin{pmatrix} 5 \\ 8 \\ -10 \\ -2 \end{pmatrix} = \begin{pmatrix} -16+9+10 \\ -16+9+16 \\ 32-18-20 \\ 64-63-4 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ -6 \\ -3 \end{pmatrix},$$

$$3[b_4] + 0[b_2] + 3[b_3] = 3\begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 2 \\ 5 \\ -8 \\ -4 \end{pmatrix} + 3\begin{pmatrix} -3 \\ 2 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 12-9 \\ 3+6 \\ -6 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ -6 \\ -3 \end{pmatrix}.$$

Первые две комбинации совпадают и дают координатный столбец вектора $x_1 \in L_1 \cap L_2$, третья и четвёртая комбинации также совпадают и дают координатный столбец вектора $x_2 \in L_1 \cap L_2$.

[6] ФСР — это линейно независимые решения однородной линейной системы, через которые линейно выражаются все решения.

Проверим линейную независимость x_1, x_2, для этого составим матрицу из их координатных столбцов: $D = \begin{pmatrix} 8 & 3 \\ 11 & 9 \\ -16 & -6 \\ -8 & -3 \end{pmatrix}$. Её минор $\begin{vmatrix} 8 & 3 \\ 11 & 9 \end{vmatrix} =$ $= 72 - 33 = 39$ является ненулевым минором порядка 2 и нет ненулевых миноров большего порядка. Этот минор является базисным минором и по теореме о базисном миноре базисные столбцы матрицы линейно независимы. Тогда сами векторы x_1, x_2 линейно независимы и образуют базис подпространства $L_1 \cap L_2$. ■

4.10. Прямая сумма подпространств. Дополнительное подпространство

Здесь рассматривается частный случай суммы подпространств линейного пространства E над полем P — прямая сумма подпространств, когда каждый вектор суммы подпространств имеет единственное разложение по подпространствам.

Определение 1.

Сумма подпространств $L = L_1 + L_2 + \ldots + L_k$ называется прямой суммой, если разложение любого вектора $x \in L$ по подпространствам $L_1, L_2, \ldots, L_k$ единственно

$$x = x_1 + x_2 + \ldots + x_k, \tag{81}$$

где $x_i \in L_i$, $i = \overline{1, k}$.

Обозначение: $L = L_1 \dot{+} L_2 \dot{+} \ldots \dot{+} L_k = \dot{\sum_{i=1}^{k}} L_i$. ■

Замечание. Мы доказали, что сумма подпространств — подпространство.

Лемма 4.10.1. *Если подпространство L является суммой подпространств L_1, L_2, …, L_k, то любой вектор $x \in L$ имеет единственное разложение (81) тогда и только тогда, когда нулевой вектор θ имеет единственное разложение $\theta = \theta + \theta + \ldots + \theta$.*

Доказательство. 1) Очевидно, что если любой вектор $x \in L$ имеет единственное разложение (81), то и нулевой вектор θ имеет единственное разложение.

2) Пусть нулевой вектор θ имеет единственное разложение. Так как L является суммой подпространств L_1, L_2, ..., L_k, то для произвольного вектора $x \in L$ существует разложение $x = x_1 + x_2 + \ldots + x_k$. Докажем, что это представление единственно. Предположим противное: существует другое разложение $x = x'_1 + x'_2 + \ldots + x'_k$, где $x'_i \in L_i$, $i = \overline{1,k}$, причём некоторая разность $x_j - x'_j \neq 0$. Вычтем $x - x = x_1 + x_2 + \ldots + x_k - x'_1 - x'_2 - \ldots - x'_k$. Отсюда получаем $\theta = (x_1 - x'_1) + (x_2 - x'_2) + \ldots + (x_k - x'_k)$, где $x_i - x'_i \in L_i$, $i = \overline{1,k}$ и $x_j - x'_j \neq 0$, а это противоречит единственности разложения θ. ■

Для двух подпространств приведём критерий прямой суммы.

Теорема 4.10.1. (Критерии прямой суммы двух подпространств)

Для ненулевых конечномерных подпространств L_1, L_2 линейного пространства E следующие условия эквивалентны.

1) *Сумма подпространств $L_1 + L_2$ — прямая.*

2) $L_1 \cap L_2 = \theta$.

3) $\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2)$.

4) *Объединение базисов подпространств L_1 и L_2 является базисом суммы подпространств.*

Доказательство. 1) $\Rightarrow$ 2) Пусть сумма $L_1 + L_2$ — прямая. Возьмём произвольный $x \in L_1 \cap L_2$. Тогда $x \in L_1$, $x \in L_2$ и для него напишем два разложения $x = x + \theta$, $x \in L_1$, $\theta \in L_2$, $x = \theta + x$, $\theta \in L_1$, $x \in L_2$. В силу единственности разложения получаем $x = \theta$, т. е. $L_1 \cap L_2 = \theta$.

2) $\Rightarrow$ 3) Если $L_1 \cap L_2 = \theta$, то по теореме 4.9.1 получаем $\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2) - \dim(L_1 \cap L_2) = \dim(L_1) + \dim(L_2)$.

3) $\Rightarrow$ 4) Из $\dim(L_1 + L_2) = \dim(L_1) + \dim(L_2)$ следует, что количество векторов в базисе подпространства $L_1 + L_2$ равно сумме количеств векторов в базисах подпространств L_1 и L_2. А по теореме 4.8.1 базисом в $L_1 + L_2$ является база объединения базисов из L_1 и L_2. Поэтому объединение базисов подпространств L_1 и L_2 является базисом суммы подпространств.

4) $\Rightarrow$ 1) Пусть $e_1, \ldots, e_k$ — базис подпространства L_1, $f_1, \ldots, f_m$ — базис подпространства L_2 и $e_1, \ldots, e_k, f_1, \ldots, f_m$ — базис суммы подпространств $L_1 + L_2$. Покажем, что нулевой вектор из $L_1 + L_2$ (а тогда в силу леммы 4.10.1 и любой вектор) имеет единственное разложение по подпространствам L_1 и L_2. Пусть $\theta = x_1 + x_2$, $x_1 \in L_1$, $x_2 \in L_2$. Разложим x_1, x_2 по базисам в L_1, L_2 соответственно: $x_1 = \sum\limits_{i=1}^{k} \alpha_i e_i$, $x_2 = \sum\limits_{i=1}^{m} \beta_i f_i$. Тогда $\theta = \sum\limits_{i=1}^{k} \alpha_i e_i + \sum\limits_{i=1}^{m} \beta_i f_i$ — разложение θ по базису в $L_1 + L_2$. У нулевого вектора все координаты α_i и β_i равны нулю, поэтому $x_1 = \sum\limits_{i=1}^{k} 0 e_i = \theta$, $x_2 = \sum\limits_{i=1}^{m} 0 f_i = \theta$. Следовательно, сумма $L_1 + L_2$ — прямая. ■

Далее с помощью дополнения базиса подпространства до базиса всего пространства строится дополнительное подпространство.

Определение 2. *Пусть L_1 подпространство пространства E. Подпространство L_2 называется дополнительным подпространством к L_1, если $L_1 \dot{+} L_2 = E$.*

При этом L_1 является дополнительным подпространством к L_2.

Простейшим примером дополнительных подпространств является пара всё пространство и нулевое подпространство.

Теорема 4.10.2. *Для любого подпространства L_1 конечномерного пространства E существует дополнительное подпространство.*

Доказательство. Поскольку $E \dot{+} \theta = E$, $\theta \dot{+} E = E$, то дополнительное для всего пространства — это нулевое подпространство, и дополнительное для нулевого подпространства — это всё пространство.

Пусть $L_1 \neq E$, $L_1 \neq \theta$ и $e_1, \ldots, e_k$ — базис L_1. Дополним эти векторы до базиса $e = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ всего пространства E. Векторы $e_{k+1}, \ldots, e_n$ линейно независимы как часть базиса и поэтому образуют базу системы $e_{k+1}, \ldots, e_n$. Обозначим $L_2 = \mathrm{Lin}(e_{k+1}, \ldots, e_n)$. Эта линейная оболочка является линейным подпространством в E с базисом $e_{k+1}, \ldots, e_n$ в силу теоремы 4.7.1. Из теоремы 4.8.1 следует, что $L_1 + L_2$ является подпространством с базисом e, тогда $\dim(L_1 + L_2) = n$. Поскольку $L_1 + L_2 \subseteq E$ и $\dim(L_1 + L_2) = n = \dim(E)$, то по теореме

4.7.2 $E = L_1 + L_2$. Для размерностей выполнены равенства $\dim(L_1) = k$, $\dim(L_2) = n - k$. Тогда $\dim(E) = \dim(L_1) + \dim(L_2)$ и из теоремы 4.10.1 следует, что сумма $E = L_1 + L_2$ — прямая. Следовательно, L_2 — дополнительное подпространство к L_1. ■

4.11. Линейное многообразие

В этом разделе в линейном пространстве E над полем P рассматриваются аналоги прямых и плоскостей из геометрии.

Определение 1. *Пусть L подпространство линейного пространства E и $a \in E$. Множество $a + L = \{a + x, x \in L\}$ называется линейным многообразием, полученным сдвигом подпространств L на вектор a. Вектор a называется вектором сдвига, а подпространство L — направляющим подпространством линейного многообразия.*

Размерностью линейного многообразия называется размерность его направляющего подпространства. ■

Теорема 4.11.1 (Свойства линейных многообразий).

1° *Если $a + L$ линейное многообразие, то выполнены следующие свойства:*

а) вектор сдвига a принадлежит $a + L$,

б) разность двух векторов из $a + L$ принадлежит L,

в) $\forall b \in a + L \quad a + L = b + L$, т. е. вектором сдвига линейного многообразия можно взять любой вектор из него.

2° *Два линейных многообразия $a + L_1$ и $b + L_2$ совпадают тогда и только тогда, когда $L_1 = L_2 = L$ и $a - b \in L$.*

Доказательство.

1° а) Поскольку $\theta \in L$, имеем $a = a + \theta \in a + L$.

б) Если $u = a + x$, $x \in L$, $v = a + y$, $y \in L$ и L подпространство, то $u - v = a + x - (a + y) = x - y \in L$.

в) Из $b \in a + L$ следует, что $b = a + x_1$, где $x_1 \in L$.

Проверим, что $a+L \subseteq b+L$.

$$\begin{cases} z \in a+L, \\ b = a+x_1,\ x_1 \in L \end{cases} \Rightarrow \begin{cases} z = a+x_2,\ x_2 \in L \\ a = b-x_1,\ x_1 \in L \end{cases} \Rightarrow \begin{cases} z = b+(-x_1+x_2), \\ (-x_1+x_2) \in L \end{cases} \Rightarrow$$
$\Rightarrow z \in b+L$.

Проверим, что $b+L \subseteq a+L$.

$$\begin{cases} z \in b+L, \\ b = a+x_1,\ x_1 \in L, \end{cases} \Rightarrow \begin{cases} z = b+x_3,\ x_3 \in L, \\ b = a+x_1,\ x_1 \in L, \end{cases} \Rightarrow \begin{cases} z = a+(x_1+x_3), \\ (x_1+x_3) \in L, \end{cases} \Rightarrow$$
$\Rightarrow z \in a+L$.

Из $a+L \subseteq b+L$ и $b+L \subseteq a+L$ следует равенство $a+L = b+L$.

2° **I.** Пусть $a+L_1 = b+L_2$, $a, b \in E$, L_1, L_2 — подпространства E.

Докажем, что $L_1 = L_2 = L$ и $a-b \in L$.

Поскольку $\theta \in L_1$, имеем $a+\theta \in a+L_1 = b+L_2$, т. е $a \in b+L_2$. Тогда из п. 1° в) следует, что $b+L_2 = a+L_2$. Использовав это равенство, условие $a+L_1 = b+L_2$ запишем в виде $a+L_1 = a+L_2$. Отсюда $\forall x \in L_1\, \exists y \in L_2$ такой, что $a+x = a+y \Rightarrow x = y \Rightarrow L_1 \subseteq L_2$, и аналогично $L_2 \subseteq L_1$. Поэтому $L_1 = L_2$. Обозначив это подпространство $L_1 = L_2 = L$, запишем условие $a \in b+L_2$ в виде $a = b+z$, где $z \in L_2 = L$. Отсюда $a-b = z \in L$.

II. Пусть $L_1 = L_2 = L$ и $a-b \in L$.

Докажем, что $a+L_1 = b+L_2$.

Обозначим $z = a-b \in L$. Тогда $a = b+z$ и $a \in b+L$. Отсюда в силу п. 1° в) следует, что $a+L = b+L$, т. е. $a+L_1 = b+L_2$. ■

4.12. Изоморфизм линейных пространств

Определение 1. *Пусть E и F — линейные пространства над общим полем P. Пространство E изоморфно пространству F, если существует биективное отображение $S\colon E \to F$ такое, что*

$$\forall x',\, x'' \in E \quad S(x'+x'') = S(x') + S(x''), \tag{82}$$

$$\forall \alpha \in P,\, x \in E \quad S(\alpha x) = \alpha S(x), \tag{83}$$

т. е. образ суммы векторов равен сумме их образов и образ произведения вектора на число равен произведению образа этого вектора на это число.

При этом отображение S называют изоморфизмом.

Обозначение: $E \cong F$. ■

Теорема 4.12.1 (Простейшие свойства изоморфизма).

Пусть E, F — линейные пространства над общим полем P, $E \cong F$ и $S\colon E \to F$ — изоморфизм. Тогда выполнены следующие свойства.

1° $S^{-1}\colon F \to E$ — *также изоморфизм.*

2° *Образ нулевого вектора $\theta_E \in E$ равен нулевому вектору $\theta_F \in F$:*

$S(\theta_1) = \theta_2$.

3° *Образ линейной комбинации равен линейной комбинации образов с теми же коэффициентами:*

$$\forall n \in N,\ \alpha_i \in P,\ x_i \in E,\ i = \overline{1,n} \quad S\Big(\sum_{i=1}^{n} \alpha_i x_i\Big) = \sum_{i=1}^{n} \alpha_i S(x_i). \qquad (84)$$

4° *Если векторы $x_i \in E$, $i = \overline{1,n}$ линейно зависимы (независимы), то их образы $S(x_i) \in F$, $i = \overline{1,n}$ линейно зависимы (независимы).*

5° *Если векторы e_i, $i = \overline{1,n}$ образуют базис в E, то их образы $S(e_i)$, $i = \overline{1,n}$ образуют базис в F.*

Доказательство.

1° 1) Если $S\colon E \to F$ — биекция, то существует обратное отображение $S^{-1}\colon F \to E$ также биективное.

2) Покажем, что $\forall y', y'' \in F \quad S^{-1}(y' + y'') = S^{-1}(y') + S^{-1}(y'')$.

Обозначим $x' = S^{-1}(y')$, $x'' = S^{-1}(y'')$, $x', x'' \in E$. Отсюда $S(x') = y'$, $S(x'') = y''$. Для изоморфизма S имеем $S(x' + x'') = S(x') + S(x'')$, тогда $S(x'+x'') = y'+y''$. Отсюда $S^{-1}(y'+y'') = x'+x''$, т. е. $S^{-1}(y'+y'') = S^{-1}(y') + S^{-1}(y'')$.

3) Покажем, что $\forall \alpha \in P, y \in F \quad S^{-1}(\alpha y) = \alpha S^{-1}(y)$.

Обозначим $x = S^{-1}(y)$, $x \in E$. Отсюда $S(x) = y$. Для изоморфизма S имеем $S(\alpha x) = \alpha S(x)$, тогда $S(\alpha x) = \alpha y$. Отсюда $S^{-1}(\alpha y) = \alpha x$,

т. е. $S^{-1}(\alpha y) = \alpha S^{-1}(y)$.

Из 1), 2), 3) следует, что $S^{-1}\colon F \to E$ — изоморфизм.

2° В равенстве $S(\alpha x) = \alpha S(x)$ возьмём $\alpha = 0$: $S(0x) = 0S(x)$. Заменим $0x = \theta_E$, $0S(x) = \theta_F$ и получим $S(\theta_E) = \theta_F$.

3° Доказываем методом математической индукции по n.

1) При $n = 1$ равенство $S(\alpha_1 x_1) = \alpha_1 S(x_1)$ следует из определения.

2) Пусть равенство верно при $n = m$: $S\Big(\sum\limits_{i=1}^{m} \alpha_i x_i\Big) = \sum\limits_{i=1}^{m} \alpha_i S(x_i)$.

3) Докажем равенство при $n = m+1$: $S\Big(\sum\limits_{i=1}^{m+1} \alpha_i x_i\Big) = \sum\limits_{i=1}^{m+1} \alpha_i S(x_i)$.

$$S\Big(\sum_{i=1}^{m+1} \alpha_i x_i\Big) = S\Big(\sum_{i=1}^{m} \alpha_i x_i + \alpha_{m+1} x_{m+1}\Big) = S\Big(\sum_{i=1}^{m} \alpha_i x_i\Big) + S(\alpha_{m+1} x_{m+1}) =$$
$$= \sum_{i=1}^{m} \alpha_i S(x_i) + \alpha_{m+1} S(x_{m+1}) = \sum_{i=1}^{m+1} \alpha_i S(x_i), \text{ ч. т. д.}$$

4° Пусть векторы $x_i \in E$, $i = \overline{1,n}$ линейно зависимы. Тогда существуют числа $\alpha_i \in P$ не все равные нулю такие, что $\sum\limits_{i=1}^{n} \alpha_i x_i = \theta_1$. Взяв образы левой и правой части, получим $S\big(\sum\limits_{i=1}^{n} \alpha_i x_i\big) = S(\theta_1)$. Использовав утверждения 3° и 2°, получим $\sum\limits_{i=1}^{n} \alpha_i S(x_i) = \theta_2$. Здесь нетривиальная линейная комбинация векторов $S(x_i)$ равна θ_2, поэтому векторы $S(x_i)$ — линейно зависимы.

Пусть векторы $x_i \in E$, $i = \overline{1,n}$ линейно независимы. $\sqcap$ Векторы $S(x_i), i = \overline{1,n}$ линейно зависимы в F. В силу 1° S^{-1} — изоморфизм, тогда по доказанному получим $S^{-1}(S(x_i))$, $i = \overline{1,n}$ линейно зависимы. Но $S^{-1}(S(x_i)) = x_i$, $i = \overline{1,n}$, тогда x_i, $i = \overline{1,n}$ линейно зависимы. $\sqcup$ Следовательно, векторы $S(x_i), i = \overline{1,n}$ линейно независимы.

5° а) Пусть векторы e_i, $i = \overline{1,n}$ образуют базис в E, тогда они линейно независимы. По утверждению 4° векторы $f_i = S(e_i)$, $i = \overline{1,n}$ линейно независимы в пространстве F. б) Покажем, что любой вектор $y \in F$ линейно выражается через векторы f_i, $i = \overline{1,n}$. Возьмём век-

тор $x = S^{-1}(y) \in E$ и разложим его по базису e_i, $i = \overline{1,n}$, получим $x = \sum\limits_{i=1}^{n} \alpha_i e_i$. Взяв образы левой и правой части, получим $S(x) = S\big(\sum\limits_{i=1}^{n} \alpha_i e_i\big) \Leftrightarrow S(x) = \sum\limits_{i=1}^{n} \alpha_i S(e_i) \Leftrightarrow y = \sum\limits_{i=1}^{n} \alpha_i f_i$.

Из а) и б) следует, что $f_i = S(e_i)$, $i = \overline{1,n}$ образуют базис в F. ■

Теорема 4.12.2 (Критерий изоморфности пространств).

Конечномерные пространства над общим полем изоморфны тогда и только тогда, когда их размерности равны.

Доказательство. Пусть E, F — линейные пространства над полем P.

I. Пусть $E \cong F$ и $S\colon E \to F$ — изоморфизм. а) Если $E = \theta_E$, то $F = S(E) = \theta_F$. Тогда $\dim(E) = \dim(F) = 0$. б) Если $E \neq \theta_E$, то $\dim(E) = n \in N$. Следовательно, базис в E состоит из n векторов. По теореме 4.12.1 базис в F состоит из n векторов, тогда $\dim(F) = n = \dim(E)$.

II. Пусть $\dim(E) = \dim(F) = n$.

а) При $n = 0$ имеем $E = \theta_E$, $F = \theta_F$. Определим отображение $S\colon E \to F$ формулой $S(\theta_E) = \theta_F$. Тогда S — биективное отображение, выполнены свойства (82), (83), поэтому $E \cong F$ и $S\colon E \to F$ — изоморфизм.

б) При $n \in N$ существуют базисы $e_1, \ldots, e_n$ и $f_1, \ldots, f_n$ в пространствах E и F соответственно.

Построим отображение $S\colon E \to F$, поставив в соответствие каждому вектору $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ вектор $y = \sum\limits_{i=1}^{n} \alpha_i f_i \in F$, который имеет те же координаты α_i, $i = \overline{1,n}$, что и вектор x.

1) В силу единственности разложения вектора по базису разным векторам $x' = \sum\limits_{i=1}^{n} \alpha'_i e_i \neq x'' = \sum\limits_{i=1}^{n} \alpha''_i e_i \Leftrightarrow \exists j : \alpha'_j \neq \alpha''_j$ соответствуют разные образы векторов $S(x') = \sum\limits_{i=1}^{n} \alpha'_i f_i \neq S(x'') = \sum\limits_{i=1}^{n} \alpha''_i f_i$. У каждого вектора $y = \sum\limits_{i=1}^{n} \alpha_i f_i \in F$ имеется прообраз $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$. Отсюда следует биективность отображения S.

Проверим свойства (82), (83).

2) Пусть $x' = \sum\limits_{i=1}^{n} \alpha'_i e_i \in E$, $x'' = \sum\limits_{i=1}^{n} \alpha''_i e_i \in E \Rightarrow x' + x'' = \sum\limits_{i=1}^{n} (\alpha'_i + \alpha''_i) e_i$, $S(x') = \sum\limits_{i=1}^{n} \alpha'_i f_i$, $S(x'') = \sum\limits_{i=1}^{n} \alpha''_i f_i$, $S(x' + x'') = \sum\limits_{i=1}^{n} (\alpha'_i + \alpha''_i) f_i$. Сложив равенства, содержащие $S(x')$ и $S(x'')$, получим равенство, содержащее $S(x' + x'')$, т. е. $S(x' + x'') = S(x') + S(x'')$.

3) Пусть $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$, $\beta \in P \Rightarrow \beta x = \sum\limits_{i=1}^{n} (\beta \alpha_i) e_i \in E$, $S(x) = \sum\limits_{i=1}^{n} \alpha_i f_i$, $S(\beta x) = \sum\limits_{i=1}^{n} (\beta \alpha_i) f_i$. Умножив равенство, содержащее $S(x)$, на число β, получим равенство, содержащее $S(\beta x)$. Отсюда $S(\beta x) = \beta S(x)$.

Из 1), 2), 3) следует, что $E \cong F$ и $S \colon E \to F$ — изоморфизм. ■

5. Линейные операторы

5.1. Линейный оператор. Свойства линейного оператора

Определение 1. *Пусть E, F линейные пространства над общим полем P. Отображение $A\colon E \to F$ называется линейным оператором (линейным отображением), действующим из пространства E в пространство F, если для любых $\alpha \in P$, x, $x_1, x_2 \in E$*

1) $A(x_1 + x_2) = A(x_1) + A(x_2)$ (образ суммы векторов равен сумме их образов);

2) $A(\alpha x) = \alpha A(x)$ (образ произведения вектора на число равен произведению образа вектора на это число).

При $F = E$ линейный оператор называется линейным преобразованием пространства E.

При $F = P$ линейный оператор называется линейным функционалом.

Множество всех линейных операторов, действующих из пространства E в пространство F, обозначается $L(E, F)$.

Операторы $A, B \in L(E, F)$ равны, если $A(x) = B(x)$ при всех $x \in E$. ■

Теорема 5.1.1.

Пусть линейный оператор $A \in L(E, F)$.

1° Образом нулевого вектора является нулевой вектор: $A(\theta_E) = \theta_F$, где θ_E, θ_F – нулевые векторы в E и F соответственно.

2° Образ линейной комбинации равен линейной комбинации образов с теми же коэффициентами, т.е. $\forall k \in N$, $\alpha_i \in P$, $x_i \in E$, $i = \overline{1, k}$ выполнено равенство:

$$A\Big(\sum_{i=1}^{k} \alpha_i x_i\Big) = \sum_{i=1}^{k} \alpha_i A(x_i).$$

3° Линейный оператор сохраняет линейную зависимость, т.е. если векторы $x_1, \ldots, x_k$ линейно зависимы, то векторы $A(x_1), \ldots, A(x_k)$ линейно зависимы.

Доказательство. 1° В определении возьмём число $\alpha = 0$, получим $A(0x) = 0A(x)$. Здесь $0x = \theta_E$, $0A(x) = \theta_F$, т. е. $A(\theta_E) = \theta_F$.

2° Доказывается по индукции.

При $k = 1$ утверждение $A(\alpha_1 x_1) = \alpha_1 A(x_1)$ верно по определению.

Пусть утверждение верно при $k = m$, т. е. $A\Big(\sum\limits_{i=1}^{m} \alpha_i x_i\Big) = \sum\limits_{i=1}^{m} \alpha_i A(x_i)$.

Докажем, что утверждение верно при $k = m + 1$:

$A\Big(\sum\limits_{i=1}^{m+1} \alpha_i x_i\Big) = \sum\limits_{i=1}^{m+1} \alpha_i A(x_i)$.

$$A\Big(\sum_{i=1}^{m+1} \alpha_i x_i\Big) \boxed{\begin{array}{l}\text{разобъём}\\ \text{сумму на}\\ \text{две части}\end{array}} = A\Big(\sum_{i=1}^{m} \alpha_i x_i + \alpha_{m+1} x_{m+1}\Big) \boxed{\begin{array}{l}\text{по опреде-}\\ \text{лению}\end{array}} =$$

$$= A\Big(\sum_{i=1}^{m} \alpha_i x_i\Big) + A(\alpha_{m+1} x_{m+1}) \boxed{\begin{array}{l}\text{по предпо-}\\ \text{ложению}\\ \text{индукции и}\\ \text{определению}\end{array}} = \sum_{i=1}^{m} \alpha_i A(x_i) + \alpha_{m+1} A(x_{m+1}) =$$

$$\boxed{\text{сложим}} = \sum_{i=1}^{m+1} \alpha_i A(x_i).$$

3° Если $x_1, \ldots, x_k$ линейно зависимы, то найдётся нетривиальная линейная комбинация этих векторов, равная нулевому вектору: $\sum\limits_{i=1}^{k} \alpha_i x_i = \theta_E$. Применим оператор A к обеим частям равенства: $A\Big(\sum\limits_{i=1}^{k} \alpha_i x_i\Big) = A(\theta_E)$. Тогда из 1° и 2° получаем $\sum\limits_{i=1}^{k} \alpha_i A(x_i) = \theta_F$, т.е. нетривиальная линейная комбинация векторов $A(e_1), \ldots, A(e_k)$ равна нулевому вектору. Следовательно, векторы $A(e_1), \ldots, A(e_k)$ линейно зависимы.

Следствие. Если векторы $A(e_1), \ldots, A(e_k)$ линейно независимы, то векторы $x_1, \ldots, x_k$ линейно независимы. ■

Свойство 2° говорит о том, что для нахождения значения линейного оператора $A \in L(E, F)$ на любом векторе достаточно знать его значения на векторах базиса $e_1, e_2, \ldots, e_n$ пространства E, поскольку образ любого вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ можно найти по формуле $A(x) = \sum\limits_{i=1}^{n} \alpha_i A(e_i)$. Реализуем эту идею в следующем утверждении.

Теорема 5.1.2 (Задание линейного оператора).

Пусть $e_1, e_2, \ldots, e_n$ — базис пространства E, и $f_1, f_2, \ldots, f_n$ — любые векторы пространства F. Тогда существует, и притом един-

ственный, линейный оператор $A \in L(E, F)$, *который переводит векторы* $e_1, e_2, \ldots, e_n$ *в векторы* $f_1, f_2, \ldots, f_n$ *соответственно, т.е.*

$$A(e_j) = f_j,\ j = \overline{1, n}. \tag{85}$$

При этом образом произвольного вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$, $(\alpha_i \in P,\ i = \overline{1, n})$ *является вектор*

$$A(x) = \sum_{i=1}^{n} \alpha_i f_i \in F. \tag{86}$$

Доказательство. I. Построим искомый оператор A, положив образом произвольного вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ вектор $\sum\limits_{i=1}^{n} \alpha_i f_i \in F$. Из единственности разложения вектора $x = \sum\limits_{i=1}^{n} \alpha_i e_i$ по базису следует, что правило (86) $\sum\limits_{i=1}^{n} \alpha_i e_i \mapsto \sum\limits_{i=1}^{n} \alpha_i f_i$ однозначно определяет образ вектора x.

II. Проверим, что для оператора A выполнены условия (85). Для вектора базиса $x = e_j$ имеем j-я координата $\alpha_j = 1$, остальные координаты $\alpha_i = 0$, $i \neq j$. Тогда из (86) получаем $A(e_j) = f_j$.

III. Проверим линейность оператора A.

1) Пусть $x_1 = \sum\limits_{i=1}^{n} \alpha'_i e_i$, $x_2 = \sum\limits_{i=1}^{n} \alpha''_i e_i$, тогда $x_1 + x_2 = \sum\limits_{i=1}^{n} (\alpha'_i + \alpha''_i) e_i$.
Отсюда по формуле (86) получаем

$$A(x_1) = \sum_{i=1}^{n} \alpha'_i f_i,\ \ A(x_2) = \sum_{i=1}^{n} \alpha''_i f_i,\ \ A(x_1 + x_2) = \sum_{i=1}^{n} (\alpha'_i + \alpha''_i) f_i.$$

Сложив $A(x_1)$ и $A(x_2)$, находим

$$A(x_1) + A(x_2) = \sum_{i=1}^{n} \alpha'_i f_i + \sum_{i=1}^{n} \alpha''_i f_i = \sum_{i=1}^{n} (\alpha'_i + \alpha''_i) f_i.$$

Сравнив два выражения, получим

$$A(x_1 + x_2) = \sum_{i=1}^{n} (\alpha'_i + \alpha''_i) f_i = A(x_1) + A(x_2), \text{ т.е. } A(x_1 + x_2) = A(x_1) + A(x_2).$$

2) Пусть $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$, $\beta \in P$, тогда $\beta x = \sum\limits_{i=1}^{n} (\beta \alpha_i) e_i$.
Отсюда по формуле (86) находим

$$A(x) = \sum_{i=1}^{n} \alpha_i f_i,\ \ A(\beta x) = \sum_{i=1}^{n} (\beta \alpha_i) f_i.$$

Умножив обе части первого равенства на число β, получим $\beta A(x) = \beta \sum\limits_{i=1}^{n} \alpha_i f_i$. Тогда, в силу $A(\beta x) = \sum\limits_{i=1}^{n} (\beta\alpha_i) f_i$, выполнено равенство $A(\beta x) = \beta A(x)$.

IV. Оператор A единственен.

Действительно, если для оператора $B \in L(E, F)$ выполнены условия $B(e_j) = f_j$, $j = \overline{1, n}$, то для произвольного $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E$ имеем

$$B(x) = B\big(\sum_{i=1}^{n} \alpha_i e_i\big) = \sum_{i=1}^{n} \alpha_i B(e_i) = \sum_{i=1}^{n} \alpha_i f_i = A(x).$$

Следовательно, $B = A$. ■

5.2. Матрица линейного оператора. Линейное выражение координат образа вектора при линейном отображении. Связь матриц линейного оператора в различных базисах

В этом разделе рассматриваем линейные пространства E, F над общим полем P и множество линейных операторов $L(E, F)$.

Определение 1. *Пусть E, F — линейные пространства с базисами $e = (e_1, e_2, \ldots, e_n)$, $f = (f_1, f_2, \ldots, f_m)$ соответственно, и A — линейный оператор, действующий из E в F.*

Разложим образы векторов базиса e по базису f :

$$A(e_j) = \sum_{i=1}^{m} a_{ij} f_i, \ j = \overline{1, n} \tag{87}$$

или подробнее

$$\begin{aligned} A(e_1) &= a_{11} f_1 + a_{21} f_2 + \ldots + a_{m1} f_m, \\ A(e_2) &= a_{12} f_1 + a_{22} f_2 + \ldots + a_{m2} f_m, \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \ldots \quad \ldots \\ A(e_n) &= a_{1n} f_1 + a_{2n} f_2 + \ldots + a_{mn} f_m. \end{aligned}$$

Из координатных столбцов векторов $[A(e_1)]_f, [A(e_2)]_f, \ldots, [A(e_n)]_f$ в базисе f составим матрицу

$$[A]_{f,e} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{88}$$

Матрица $[A]_{f,e}$ *называется матрицей линейного оператора* A *в паре базисов* e *и* f. ■

Замечание. Из единственности разложения вектора по базису следует, что при фиксированных базисах матрица линейного оператора определена однозначно. Итак, каждому оператору $A \in L(E, F)$ соответствует его матрица $[A]_{f,e} \in P^{m \times n}$.

Таким образом, при фиксированных базисах определено отображение, действующее из множества линейных операторов $L(E, F)$ во множество матриц $P^{m \times n}$ по правилу

$$A \to [A]_{e,f}. \tag{89}$$

Теорема 5.2.1.
Отображение (89) биективно.

Доказательство. Инъективность. Пусть A, B линейные операторы из $L(A, B)$ и $A \neq B$. В силу теоремы 5.1.2 значения $A(e_i)$ и $B(e_i)$ на всех базисных векторах e_i не могут совпадать, иначе если $A(e_i) = B(e_i)$ при всех $i = \overline{1, n}$, то $A = B$. Поэтому координатные столбцы несовпадающих векторов $A(e_i) \neq B(e_i)$ различны, и поэтому матрицы $[A]_{f,e} \neq [B]_{f,e}$.

Сюръективность. Покажем, что любая матрица $H = (h_{ij}) \in P^{m \times n}$ является матрицей некоторого линейного оператора из $L(E, F)$. Возьмём векторы $g_j = \sum\limits_{i=1}^{m} h_{ij} f_i,\ j = \overline{1, n}$ и согласно теореме 5.1.2 построим линейный оператор $C \in L(E, F)$, для которого $C(e_j) = g_j,\ j = \overline{1, n}$. Для него $C(e_j) = g_j = \sum\limits_{i=1}^{m} h_{ij} f_i,\ j = \overline{1, n}$ и из определения 1 следует, что матрица $[C]_{f,e}$ совпадает с матрицей H. ■

Установим связь координат вектора и его образа при линейном отображении.

Теорема 5.2.2 (Связь координат вектора и его образа).

Пусть E, F — линейные пространства с базисами $e = (e_1, e_2, \ldots$ $\ldots, e_n)$, $f = (f_1, f_2, \ldots, f_m)$ соответственно, A — линейный оператор, действующий из E в F, $x \in E$, $y \in F$.

Тогда следующие условия эквивалентны:

$$y = A(x); \tag{90}$$

$$[y]_f = [A]_{f,e}[x]_e; \tag{91}$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}, \tag{92}$$

где $[x]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$, $[y]_f = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_m \end{pmatrix}$ *— координатные столбцы вектора x в базисе e и вектора $y = A(x)$ в базисе f соответственно, $A_{f,e} =$ $= (a_{ij})$ — матрица оператора A в паре базисов e, f.*

Доказательство. 1) Формула (91) — это краткая запись формулы (92).

2) Пусть верна формула (90): $y = A(x)$, где $x = \sum\limits_{j=1}^{n} \alpha_j e_j$, $y =$ $= \sum\limits_{i=1}^{m} \beta_i f_i$.

Тогда $y = A(x) = A\big(\sum\limits_{j=1}^{n} \alpha_j e_j\big) = \sum\limits_{j=1}^{n} \alpha_j A(e_j) \boxed{\text{подставим (87)}} = \sum\limits_{j=1}^{n} \alpha_j \sum\limits_{i=1}^{m} a_{ij} f_i \boxed{\text{внесём } \alpha_j} =$

$= \sum\limits_{j=1}^{n} \sum\limits_{i=1}^{m} a_{ij} \alpha_j f_i = \boxed{\text{поменем порядок суммирования}} = \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} a_{ij} \alpha_j f_i \boxed{\text{вынесем } f_i} = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij} \alpha_j\big) f_i.$

Итак, $y = \sum\limits_{i=1}^{m} \beta_i f_i$ и $y = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij} \alpha_j\big) f_i$. Отсюда в силу единственности разложения по базису имеем $\beta_i = \sum\limits_{j=1}^{n} a_{ij} \alpha_j$, $i = \overline{1, n}$, а эти формулы в матричной записи дают (92).

3) Пусть верна формула (92). Из неё следует $\beta_i = \sum\limits_{j=1}^{n} a_{ij}\alpha_j$, $i = \overline{1,m}$. Подставим в вектор $y = \sum\limits_{i=1}^{m} \beta_i e_i$ выражения координат $\beta_i = \sum\limits_{j=1}^{n} a_{ij}\alpha_j$, получим $y = \sum\limits_{i=1}^{m} \big(\sum\limits_{j=1}^{n} a_{ij}\alpha_j \big) f_i$. Выполнив преобразования п. 2) в обратном порядке, приходим к выводу, что $y = A(x)$. ■

Выясним связь матриц линейного оператора в различных базисах.

Теорема 5.2.3 (Связь матриц оператора в различных базисах).

Пусть в пространстве E заданы два базиса $e = (e_1, e_2, \ldots, e_n)$, $e' = (e'_1, e'_2, \ldots, e'_n)$ с матрицей перехода $T_{e\to e'}$, в пространстве F заданы два базиса $f = (f_1, f_2, \ldots, f_m)$, $f' = (f'_1, f'_2, \ldots, f'_m)$ с матрицей перехода $T_{f\to f'}$, линейному оператору $A \in L(E,F)$ в паре базисов e, f соответствует матрица $[A]_{f,e}$, а в паре базисов e', f' — матрица $[A]_{f',e'}$.

Тогда матрицы $[A]_{f,e}$ и $[A]_{f',e'}$ связаны соотношениями

$$T_{f\to f'}[A]_{f',e'} = [A]_{f,e}T_{e\to e'}, \tag{93}$$

$$[A]_{f',e'} = T^{-1}_{f\to f'}[A]_{f,e}T_{e\to e'}. \tag{94}$$

Доказательство.

Для вектора $x \in E$ и его образа $y = A(x)$ в силу (91) имеем

$$[y]_f = [A]_{f,e}[x]_e, . \tag{95}$$

$$[y]_{f'} = [A]_{f',e'}[x]_{e'}. \tag{96}$$

В свою очередь, согласно (74), $[x]_e = T_{e\to e'}[x]_{e'}$, $[y]_f = T_{f\to f'}[y]_{f'}$. Подставив эти соотношения в (95), получим

$$T_{f\to f'}[y]_{f'} = [A]_{f,e}T_{e\to e'}[x]_{e'}.$$

Далее подставив сюда (96), получим

$$T_{f\to f'}[A]_{f',e'}[x]_{e'} = [A]_{f,e}T_{e\to e'}[x]_{e'}. \quad \big(T_{f\to f'}[A]_{f',e'}\big)[x]_{e'} = \big([A]_{f,e}T_{e\to e'}\big)[x]_{e'}.$$

Поскольку последнее равенство выполняется для любых координатных столбцов $[x]_{e'}$, то выполнено равенство $T_{f\to f'}[A]_{f',e'} = [A]_{f,e}T_{e\to e'}$. Откуда после умножения на обратную матрицу $T^{-1}_{f\to f'}$ получаем (94). ■

Замечание. *Для линейного преобразования* $A\colon E \to E$ *при построении его матрицы часто берут один базис* $e = f = (e_1, e_2, \ldots, e_n)$, *находят матрицу* $[A]_{e,e}$, *и называют её матрицей линейного преобразования* A *в базисе* e.

Тогда справедливо следующее утверждение.

Следствие. *Пусть в линейном пространстве* E *заданы два базиса* $e = (e_1, e_2, \ldots, e_n)$ *(«старый») и* $e' = (e'_1, e'_2, \ldots, e'_n)$ *(«новый») с матрицей перехода* $T_{e\to e'}$, *линейному преобразованию* $A \in L(E, E)$ *в базисе* e *соответствует матрица* $[A]_{e,e}$, *а в базисе* e' *— матрица* $[A]_{e',e'}$.

Тогда матрицы $[A]_{e,e}$ *и* $[A]_{e',e'}$ *связаны соотношениями*

$$T_{e\to e'}[A]_{e',e'} = [A]_{e,e}T_{e\to e'}, \tag{97}$$

$$[A]_{e',e'} = T^{-1}_{e\to e'}[A]_{e,e}T_{e\to e'}. \tag{98}$$

5.3. Сумма операторов, произведение оператора на число, произведение операторов и соответствующие им матрицы

В этом разделе введём операции сложения линейных операторов, умножения линейных операторов на число, умножения операторов, покажем, что результирующие операторы являются линейными операторами и найдём их матрицы.

Определение 1.

Пусть E, F, G *— линейные пространства над полем* P.

а) *Суммой линейных операторов* $A, B \in L(E, F)$ *называется отображение из* E *в* F, *выполняемое по правилу* $x \mapsto A(x) + B(x)$, $\forall x \in E$. *Обозначение:* $A + B$. *Итак,*

$$(A + B)(x) = A(x) + B(x),\ \forall x \in E. \tag{99}$$

б) *Произведением линейного оператора* $A \in L(E, F)$ *на число* $\alpha \in P$ *называется отображение из* E *в* F, *выполняемое по правилу* $x \mapsto$ $\mapsto \alpha A(x)\, \forall x \in E$. *Обозначение:* αA. *Итак,*

$$(\alpha A)(x) = \alpha A(x),\ \forall x \in E. \tag{100}$$

в) *Произведением линейного оператора $A \in L(E,F)$ на линейный оператор $B \in L(F,G)$ называется отображение из E в G, выполняемое по правилу $x \mapsto B(A(x))$ $\forall x \in E$. Обозначение BA. Итак,*

$$(BA)(x) = B(A(x)),\ \forall x \in E. \blacksquare \tag{101}$$

Теорема 5.3.1.

1° *Для любых линейных операторов $A, B \in L(E,F)$ $A + B \in$ $\in L(E,F)$.*

2° *Для любого линейного оператора $A \in L(E,F)$ и числа $\alpha \in P$ $\alpha A \in L(E,F)$.*

3° *Для любых линейных операторов $A \in L(E,F)$, $B \in L(F,G)$ $BA \in L(E,G)$.*

Доказательство.

1° а) $\forall x_1, x_2 \in E \quad (A+B)(x_1+x_2) = (A+B)(x_1) + (A+B)(x_2)$.

$(A+B)(x_1+x_2) \boxed{\text{по определению суммы линейных операторов}} = A(x_1+x_2) + B(x_1+x_2) =$

$\boxed{\text{по определению линейных операторов}} = A(x_1) + A(x_2) + B(x_1) + B(x_2) \boxed{\text{изменим порядок слагаемых}} =$

$= A(x_1) + B(x_1) + A(x_2) + B(x_2) \boxed{\text{по определению суммы линейных операторов}} = (A+B)(x_1) +$
$+ (A+B)(x_2)$.

б) $\forall x \in E,\ \alpha \in P \quad (A+B)(\alpha x) = \alpha(A+B)(x)$.

$(A+B)(\alpha x) \boxed{\text{по определению суммы линейных операторов}} = A(\alpha x) + B(\alpha x) \boxed{\text{по определению линейного оператора}} =$

$= \alpha A(x) + \alpha B(x) \boxed{\text{вынесем } \alpha \text{ за скобки}} = \alpha(A(x) + B(x)) \boxed{\text{по определению суммы линейных операторов}} =$
$= \alpha(A+B)(x)$.

2° а) $\forall x_1, x_2 \in E,\ \alpha \in P \quad (\alpha A)(x_1+x_2) = (\alpha A)(x_1) + (\alpha A)(x_2)$.

$(\alpha A)(x_1+x_2) \boxed{\text{по определению произведения оператора на число}} = \alpha A(x_1+x_2) \boxed{\text{по определению линейного оператора}} =$

$= \alpha(A(x_1) + A(x_2)) \boxed{\text{раскроем скобки}} = \alpha A(x_1) + \alpha A(x_2) \boxed{\text{по определению произведения оператора на число}} =$
$= (\alpha A)(x_1) + (\alpha A)(x_2)$.

б) $\forall x \in E,\ \alpha, \beta \in P \quad (\alpha A)(\beta x) = \beta(\alpha A)(x)$.

$(\alpha A)(\beta x) \boxed{\text{по определению произведения оператора на число}} = \alpha A(\beta x) \boxed{\text{по определению линейного оператора}} = \alpha\beta A(x) =$

$\boxed{\text{поменяем множители}} = \beta\alpha(A(x)) \boxed{\text{по определению произведения оператора на число}} = \beta(\alpha A)(x).$

3° а) $\forall x_1, x_2 \in E \quad (BA)(x_1 + x_2) = (BA)(x_1) + (BA)(x_2).$

$(BA)(x_1 + x_2) \boxed{\text{по определению произведения операторов}} = B(A(x_1 + x_2)) \boxed{\text{по определению линейного оператора}} =$

$= B(A(x_1) + A(x_2)) \boxed{\text{по определению линейного оператора}} = B(A(x_1)) + B(A(x_2)) =$

$\boxed{\text{по определению произведения операторов}} = (BA)(x_1) + (BA)(x_2).$

б) $\forall x \in E,\ \alpha \in P \quad (BA)(\alpha x) = \alpha(BA)(x).$

$(BA)(\alpha x) \boxed{\text{по определению произведения операторов}} = B(A(\alpha x)) \boxed{\text{по определению линейного оператора}} = B(\alpha A(x)) =$

$\boxed{\text{по определению линейного оператора}} = \alpha B(A(x)) \boxed{\text{по определению произведения операторов}} = \alpha(BA)(x).$ ■

Теорема 5.3.2.

Пусть E, F линейные пространства над полем P. Тогда множество линейных операторов $L(E, F)$ с операциями сложения операторов и умножения операторов на числа из P является линейным пространством.

Доказательство. Все операторные равенства проверяются на произвольных векторах линейного пространства, причём получаемые равенства векторов выполнены.

1. $\forall A,\ B, C \in L(E, F)\ (A+B)+C = A+(B+C)$ (ассоциативность сложения операторов).

$(A + B) + C = A + (B + C) \Leftrightarrow \forall x \in E\ ((A + B) + C)(x) = (A + (B + C))(x) \Leftrightarrow \forall x \in E\ (A + B)(x) + C(x) = A(x) + (B + C))(x) \Leftrightarrow$
$\Leftrightarrow \forall x \in E\ (A(x) + B(x)) + C(x) = A(x) + (B(x) + C(x))$. Последнее равенство верно, т. к. сложение векторов ассоциативно.

2. Нулевой элемент пространства $L(E, F)$ — это нулевое отображение O, выполняемое по правилу $\forall x \in E\ O(x) = \theta_F$, где θ_F — нулевой вектор в F.

Для $A \in L(E, F)$ проверим $A + O = A$.

$A + O = A \Leftrightarrow \forall x \in E \ (A + O)(x) = A(x) \Leftrightarrow \forall x \in E \ A(x) + O(x) = A(x) \Leftrightarrow \forall x \in E \ A(x) + \theta_F = A(x)$. Последнее равенство верно, т. к. θ_F — нулевой вектор в F.

3. Противоположный элемент к оператору $A \in L(E, F)$ — это отображение $(-A)$, выполняемое по правилу $\forall x \in E \ (-A)(x) = -A(x)$.

Для $A \in L(E, F)$ проверим $A + (-A) = O$.
$A + (-A) = O \Leftrightarrow \forall x \in E \ (A + (-A))(x) = O(x) \Leftrightarrow \forall x \in E \ A(x) + (-A)(x) = \theta_F \Leftrightarrow \forall x \in E \ A(x) + (-A(x)) = \theta_F \Leftrightarrow \theta_F = \theta_F$. Последнее равенство верно.

4. $\forall A, B \in L(E, F) \ A + B = B + A$ (коммутативность сложения операторов).
$A + B = B + A \Leftrightarrow \forall x \in E \ (A + B)(x) = (B + A)(x) \Leftrightarrow \forall x \in E \ A(x) + B(x) = B(x) + A(x)$. Последнее равенство верно, т. к. сложение векторов коммутативно.

5. $\forall \alpha, \beta \in P, A \in L(E, F) \ \alpha(\beta A) = (\alpha\beta)A$.
$\alpha(\beta A) = (\alpha\beta)A \Leftrightarrow \forall x \in E \ (\alpha(\beta A))(x) = ((\alpha\beta)A)(x) \Leftrightarrow$
$\Leftrightarrow \forall x \in E \ \alpha(\beta A)(x) = (\alpha\beta)A(x) \Leftrightarrow \forall x \in E \ \alpha(\beta(A(x)) = (\alpha\beta)A(x)$. Последнее равенство верно, т. к. верна аксиома АЛП5.

Остальные аксиомы проверьте самостоятельно. ■

Теорема 5.3.3. 1° *Пусть в линейных пространствах E, F над полем P фиксированы базисы $e = (e_1, \ldots, e_n)$ и $f = (f_1, \ldots, f_m)$ соответственно. Тогда*

а) для любых линейных операторов A, $B \in L(E, F)$ в паре базисов e, f матрица суммы операторов равна сумме их матриц:

$$[A + B]_{f,e} = [A]_{f,e} + [B]_{f,e}\,, \tag{102}$$

б) для любого линейного оператора $A \in L(E, F)$ и числа $\alpha \in P$ в паре базисов e, f матрица произведения оператора на число равна произведению матрицы оператора на это число:

$$[\alpha A]_{f,e} = \alpha[A]_{f,e}. \tag{103}$$

2° *Пусть в линейных пространствах E, F, G над полем P заданы базисы $e = (e_1, \ldots, e_n)$, $f = (f_1, \ldots, f_m)$, $g = (g_1, \ldots, g_k)$ соответственно и определены линейные операторы $A \in L(E, F)$ и $B \in L(F, G)$.*

Тогда матрица произведения этих операторов равна произведению матриц этих операторов в соответствующих парах базисов:

$$[BA]_{g,e} = [B]_{g,f}[A]_{f,e}. \tag{104}$$

Доказательство.

1° а) Пусть $[A]_{f,e} = (a_{ij})$, $[B]_{f,e} = (b_{ij})$. Из определения матриц операторов имеем $A(e_j) = \sum\limits_{i=1}^{m} a_{ij} f_i$, $B(e_j) = \sum\limits_{i=1}^{m} b_{ij} f_i$. Поэтому $(A+B)(e_j) = A(e_j) + B(e_j) = \sum\limits_{i=1}^{m} (a_{ij} + b_{ij}) f_i$. Отсюда матрица оператора $A+B$ равна $(a_{ij} + b_{ij})$, что совпадает с суммой матриц $(a_{ij}) + (b_{ij})$.

1° б) Пусть $[A]_{f,e} = (a_{ij})$. Тогда $A(e_j) = \sum\limits_{i=1}^{m} a_{ij} f_i$ и $(\alpha A)(e_j) = \alpha A(e_j) = \alpha \sum\limits_{i=1}^{m} a_{ij} f_i = \sum\limits_{i=1}^{m} (\alpha a_{ij}) f_i$. Отсюда матрица оператора αA равна (αa_{ij}), что совпадает с произведением $\alpha(a_{ij})$.

2° Пусть $[A]_{f,e} = (a_{ij})$, $[B]_{g,f} = (b_{ij})$, $[BA]_{g,e} = (c_{ij})$. Составим образы векторов $(BA)(e_j)$ двумя способами: 1) $(BA)(e_j) = \sum\limits_{i=1}^{k} c_{ij} g_i$,

2) $(BA)(e_j) = B(A(e_j)) = B\Big(\sum\limits_{s=1}^{m} a_{sj} f_s\Big) = \sum\limits_{s=1}^{m} a_{sj} B(f_s) = \sum\limits_{s=1}^{m} a_{sj} \sum\limits_{i=1}^{k} b_{is} g_i = \sum\limits_{s=1}^{m} \sum\limits_{i=1}^{k} a_{sj} b_{is} g_i = \sum\limits_{i=1}^{k} \Big(\sum\limits_{s=1}^{m} b_{is} a_{sj}\Big) g_i$. В силу единственности координат вектора приравняем координаты в двух разложениях и получим $c_{ij} = \sum\limits_{s=1}^{m} b_{is} a_{sj}$. Это означает равенство матриц $[BA]_{g,e} = [B]_{g,f}[A]_{f,e}$. ■

5.4. Образ и ядро линейного оператора

В этом разделе рассматривается некоторые множества векторов, связанных с линейным оператором.

Определение 1. *Пусть E, F — линейные пространства над полем P и A — линейный оператор из $L(E,F)$.*

Образом линейного оператора A называется множество векторов

$$\operatorname{im}(A) = \{y \in F \mid \exists x \in E : A(x) = y\}. \tag{105}$$

Ядром линейного оператора A называется множество векторов

$$\ker(\mathrm{A}) = \{x \in E \mid A(x) = \theta\}. \tag{106}$$

Теорема 5.4.1. *Если $A \in L(E, F)$, то $\ker(A)$ — линейное подпространство в E, $\operatorname{im}(A)$ — линейное подпространство в F.*

Доказательство. Используем критерий подпространства.

I. Рассмотрим $\ker(A)$. Т. к. $A(\theta) = \theta$, то $\theta \in \ker(A)$ и $\ker(A) \neq \emptyset$.

1) $x_1, x_2 \in \ker(A) \Rightarrow x_1 + x_2 \in \ker(A)$.

$x_1,\ x_2 \in \ker(A) \Rightarrow A(x_1) = \theta,\ A(x_2) = \theta \Rightarrow A(x_1) + A(x_2) = \theta \Rightarrow A(x_1 + x_2) = \theta \Rightarrow x_1 + x_2 \in \ker(A)$.

2) $x \in \ker(A),\ \alpha \in P \Rightarrow \alpha x \in \ker(A)$.

$x \in \ker(A),\ \alpha \in P \Rightarrow A(x) = \theta \Rightarrow \alpha A(x) = \theta \Rightarrow A(\alpha x) = \theta \Rightarrow \alpha x \in \ker(A)$.

I. Рассмотрим $\operatorname{im}(A)$. Поскольку $A(\theta) = \theta$, то $\operatorname{im}(A) \neq \emptyset$.

1) $y_1, y_2 \in \operatorname{im}(A) \Rightarrow y_1 + y_2 \in \operatorname{im}(A)$.

$y_1, y_2 \in \operatorname{im}(A) \Rightarrow \exists x_1,\ x_2 \in E \mid y_1 = A(x_1),\ y_2 = A(x_2) \Rightarrow y_1 + y_2 = A(x_1) + A(x_2) \Rightarrow y_1 + y_2 = A(x_1 + x_2) \Rightarrow y_1 + y_2 \in \operatorname{im}(A)$.

2) $y \in \operatorname{im}(A),\ \alpha \in P \Rightarrow \alpha y \in \operatorname{im}(A)$.

$y \in \operatorname{im}(A),\ \alpha \in P \Rightarrow \exists x \in E \mid A(x) = y \Rightarrow \alpha A(x) = \alpha y \Rightarrow A(\alpha x) = \alpha y \Rightarrow \alpha y \in \operatorname{im}(A)$. ■

Определение 2. *Рангом линейного оператора A называется размерность его образа:* $\operatorname{rg}(A) = \dim(\operatorname{im}(A))$.

Дефектом линейного оператора A называется размерность его ядра: $\operatorname{def}(A) = \dim(\ker(A))$.

Теорема 5.4.2. *Если $A \in L(E, F)$ и $e_1, \ldots, e_n$ — базис пространства E, то* $\operatorname{im}(A) = \operatorname{Lin}\big(A(e_1), \ldots, A(e_n)\big)$.

Доказательство. 1) $\operatorname{im}(A) \subseteq \operatorname{Lin}\big(A(e_1), \ldots, A(e_n)\big)$.

$y \in \operatorname{im}(A) \Rightarrow \exists x \in E \mid y = A(x) \Rightarrow \boxed{\text{разложим } x \text{ по базису } x = \sum\limits_{i=1}^{n} \alpha_i e_i} \Rightarrow y = A(x) =$

$= A\Big(\sum\limits_{i=1}^{n} \alpha_i e_i\Big) = \sum\limits_{i=1}^{n} \alpha_i A(e_i) \in \mathrm{Lin}\big(A(e_1),\, \ldots,\, A(e_n)\big).$

2) $\mathrm{Lin}\big(A(e_1),\, \ldots,\, A(e_n)\big) \subseteq \mathrm{im}(A).$

$y \in \mathrm{Lin}\big(A(e_1),\, \ldots,\, A(e_n)\big) \Rightarrow y = \sum\limits_{i=1}^{n} \alpha_i A(e_i) = A\Big(\sum\limits_{i=1}^{n} \alpha_i e_i\Big) = A(x),$ где $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in E \Rightarrow y \in \mathrm{im}(A).$ ■

Теорема 5.4.3. *Пусть E, F — конечномерные пространства и $A \in L(E, F)$. Тогда ранг оператора A равен рангу его матрицы в произвольной паре базисов.*

Доказательство. Отметим, что ранг оператора A равен размерности его образа $\mathrm{im}(A)$, а образ не зависит от выбора базиса.

Пусть $e = (e_1,\, \ldots,\, e_n)$ и $f = (f_1,\, \ldots,\, f_m)$ — произвольные базисы пространств E и F. В силу теоремы 5.4.2 $\mathrm{im}(A) = \mathrm{Lin}\big(A(e_1),\, \ldots,\, A(e_n)\big)$, а по теореме 4.7.1 базисом подпространства $\mathrm{Lin}\big(A(e_1),\, \ldots,\, A(e_n)\big)$ является база системы векторов $A(e_1),\, \ldots, A(e_n)$. Поэтому размерность этого подпространства совпадает с числом векторов базы указанной системы векторов: $\mathrm{rg}(A) = \mathrm{rg}\big(A(e_1),\, \ldots,\, A(e_n)\big)$.

Произвольная подсистема системы векторов $A(e_1),\, \ldots,\, A(e_n)$ линейно независима тогда и только тогда, когда соответствующая подсистема системы их координатных столбцов $[A(e_1)]_f,\, \ldots, [A(e_n)]_f$ линейно независима. Тогда базе системы векторов $A(e_1),\, \ldots,\, A(e_n)$ соответствуют базисные столбцы матрицы $[A]_{f,e} = \big([A(e_1)]_f\ \ldots\ [A(e_n)]_f\big)$ и ранг оператора A равен рангу его матрицы в паре базисов e, f : $\mathrm{rg}(A) = \mathrm{rg}\big([A]_{f,e}\big)$. ■

Теорема 5.4.4 (о ранге и дефекте). *Пусть E — n-мерное пространство и $A \in L(E, F)$. Тогда*

$$\mathrm{rg}(A) + \mathrm{def}(A) = n. \tag{107}$$

Доказательство. Пусть $e_1,\, \ldots,\, e_k$ базис $\ker(A)$.

1) Если $k = n$, то $\dim(\ker(A)) = n$ и $\ker(A) = E$. В этом случае $A(x) = \theta$ для любого $x \in E$, т. е. оператор нулевой и его образ $\mathrm{im}(A) = \theta$.

Тогда $\mathrm{rg}(A) = \dim(\mathrm{im}(A)) = 0$, $\mathrm{def}(A) = \dim(\ker(A)) = n$ и равенство (107) выполнено.

2) Если $k < n$, то дополним $e_1, \ldots, e_k$ до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ пространства E. Согласно теореме 5.4.2 $\mathrm{im}(A) = \mathrm{Lin}\big(A(e_1), \ldots, A(e_k), A(e_{k+1}), \ldots, A(e_n)\big)$. Здесь $A(e_1) = \ldots = A(e_k) = \theta$, т. к. векторы $e_1, \ldots, e_k$ из ядра $\ker(A)$. Поэтому $\mathrm{im}(A) = \mathrm{Lin}\big(A(e_{k+1}), \ldots, A(e_n)\big)$.

Докажем, что из равенства $\alpha_{k+1}A(e_{k+1}) + \ldots + \alpha_n A(e_n) = \theta$ следует, что все коэффициенты $\alpha_{k+1} = \ldots = \alpha_n = 0$. Запишем равенство в виде $A\big(\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n\big) = \theta$. Отсюда вектор $\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n \in \ker(A)$ и он раскладывается по базису $e_1, \ldots, e_k$: $\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n = \alpha_1 e_1 + \ldots + \alpha_k e_k$. Это приводит к равенству $-\alpha_1 e_1 - \ldots - \alpha_k e_k + \alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n = \theta$ для линейно независимых векторов базиса $e_1, \ldots, e_n$, поэтому все коэффициенты $\alpha_i = 0$, $i = \overline{1, n}$. Следовательно, векторы $A(e_{k+1}), \ldots, A(e_n)$ линейно независимы. Таким образом, $\mathrm{rg}(A) = \dim(\mathrm{im}(A)) = n - k$, $\mathrm{def}(A) = \dim(\ker(A)) = k$ и равенство (107) выполнено. ■

5.5. Собственные векторы и собственные значения линейного преобразования

В этом разделе рассматриваем линейное пространство E над полем P, вводим определение собственных векторов и собственных значений линейного преобразования в E, изучаем их свойства и способы нахождения.

Определение 1.

Ненулевой вектор $x \in E$ называется собственным вектором линейного преобразования $A \in L(E, E)$, если существует такое число $\lambda \in P$, что

$$A(x) = \lambda x. \tag{108}$$

Число λ называется собственным значением преобразования A, соответствующим собственному вектору x.

При этом говорят также, что собственный вектор x соответствует собственному значению λ.

Множество всех собственных значений преобразования A называется его спектром. Обозначение: $\sigma(A)$.

Примеры. 1. В пространстве многочленов $R[x]$ рассмотрим линейное преобразование дифференцирования $A(f) = f'$, $f \in R[x]$. Для многочлена нулевой степени $f(x) = c \neq 0$ имеем $A(c) = c' = 0$, поэтому $c' = 0c$. Итак, для преобразования дифференцирования многочлен нулевой степени является собственным вектором, соответствующим собственному значению $\lambda = 0$.

2. В пространстве функций $f(x)$ $(f\colon R \to R)$, имеющих производные всех порядков рассмотрим преобразование $A(f) = f''$ (f'' — вторая производная от функции f). Здесь любое число $\lambda \in R$ является собственным значением. Например, $\lambda = 4$, т.к. $A(e^{2x}) = (e^{2x})'' = (2e^{2x})' = 4e^{2x}$. Здесь собственный вектор (его называют собственной функцией) для $\lambda = 2$ равен e^{2x}.

3. В пространстве V_2 всех геометрических векторов плоскости, у которых начало совпадает с точкой O, рассмотрим преобразование A_α поворота векторов в плоскости вокруг точки O на угол α против часовой стрелки. При $\alpha = \pi(2k+1)$, $k \in Z$ все ненулевые векторы $\overrightarrow{x}$ являются собственными векторами, соответствующие собственному значению $\lambda = -1$, т.к. в этом случае $A_\alpha \overrightarrow{x} = -\overrightarrow{x}$. При $\alpha = 2\pi k$, $k \in Z$ есть собственные векторы и собственные значения (найдите их). При $\alpha \neq \pi k$, $k \in Z$ собственных значений и собственных векторов нет (почему?). ■

Определение 2. *Пусть $A \in L(E, E)$. Для каждого числа $\lambda \in P$ обозначим через E_λ множество всех векторов $x \in E$, удовлетворяющих уравнению* (108).

Лемма 5.5.1. *Пусть $\lambda_0 \in \sigma(A)$. Тогда выполнены утверждения:*

1° *E_{λ_0} состоит из множества всех собственных векторов преобразования A, соответствующих собственному значению λ_0, дополненному нулевым вектором.*

2° *E_{λ_0} является подпространством в пространстве E.*

Доказательство. 1° Для любого вектора x из E_{λ_0} верно равенство $A(x) = \lambda_0 x$. Отсюда следует либо $x = \theta$, либо $x \neq \theta$ и тогда x — собственный вектор, соответствующий λ_0.

2° 1) Проверим, что из $x, y \in E_{\lambda_0}$ следует $x + y \in E_{\lambda_0}$.

$x, y \in E_{\lambda_0} \Rightarrow A(x) = \lambda_0 x,\ A(y) = \lambda_0 y \ \boxed{\text{сложим}} \Rightarrow A(x) + A(y) = \lambda_0 x + \lambda_0 y \Rightarrow \boxed{\begin{array}{l}\text{из линейности}\\ A(x) + A(y) = \\ A(x+y)\end{array}} A(x+y) = \lambda_0(x+y) \Rightarrow x + y \in E_{\lambda_0}$.

2) Проверим, что из $x \in E_{\lambda_0}$, $\alpha \in P$ следует $\alpha x \in E_{\lambda_0}$.

$\begin{cases} x \in E_{\lambda_0} \\ \alpha \in P \end{cases} \Rightarrow A(x) = \lambda_0 x \ \boxed{\begin{array}{l}\text{умножим}\\ \text{на } \alpha\end{array}} \Rightarrow \alpha A(x) = \alpha\lambda_0 x \ \boxed{\begin{array}{l}\text{из линейности}\\ \alpha A(x) = A(\alpha x)\end{array}} \Rightarrow$
$\Rightarrow A(\alpha x) = \lambda_0(\alpha x) \Rightarrow \alpha x \in E_{\lambda_0}$. ■

Теорема 5.5.1.

Собственные векторы $x_1, x_2, \ldots, x_k$, соответствующие различным собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_k$, линейно независимы.

Доказательство. Применим индукцию по k.

1) Для $k = 1$ утверждение верно, так как собственный вектор является ненулевым и, следовательно, линейно независим.

2) Пусть утверждение верно для m векторов.

2) Докажем утверждение для $m + 1$ векторов $x_1, \ldots, x_m, x_{m+1}$.

Приравняем линейную комбинацию векторов нулевому вектору:

$$\alpha_1 x_1 + \ldots + \alpha_m x_m + \alpha_{m+1} x_{m+1} = \theta. \tag{109}$$

Применим преобразование A к обеим частям равенства (109)

$$A(\alpha_1 x_1 + \ldots + \alpha_m x_m + \alpha_{m+1} x_{m+1}) = A(\theta).$$

Из свойств линейного оператора, определения собственного вектора:

$A\Big(\sum\limits_{i=1}^{m+1} \alpha_i x_i\Big) = \sum\limits_{i=1}^{m+1} \alpha_i A(x_i),\ A(\theta) = \theta,\ A(x_i) = \lambda_i x_i$ следует равенство

$$\alpha_1\lambda_1 x_1 + \ldots + \alpha_m\lambda_m x_m + \alpha_{m+1}\lambda_{m+1} x_{m+1} = \theta. \tag{110}$$

Чтобы убрать из линейной комбинации последнее слагаемое, вычтем из (110) равенство (109) умноженное на λ_{m+1}, и получим

$$\alpha_1(\lambda_1 - \lambda_{m+1})x_1 + \ldots + \alpha_m(\lambda_m - \lambda_{m+1})x_m = \theta.$$

В силу предположения индукции векторы $x_1, \ldots, x_m$ линейно независимы, следовательно, все коэффициенты полученной комбинации равны нулю: $\alpha_i(\lambda_i - \lambda_{m+1}) = 0$, $i = \overline{1, m}$. Отсюда, поскольку все собственные значения различны, имеем $(\lambda_i - \lambda_{m+1}) \neq 0$, и тогда $\alpha_i = 0$, $i = \overline{1, m}$. При этом из (109) следует равенство $\alpha_{m+1} x_{m+1} = \theta$, и, поскольку $x_{m+1} \neq \theta$, получаем $\alpha_{m+1} = 0$. Следовательно, в линейной комбинации (109) все коэффициенты α_i равны 0 и векторы $x_1, \ldots, x_m, x_{m+1}$ линейно независимы. ■

Следствия. 1. *Линейное преобразование n-мерного пространства не может иметь более чем n различных собственных значений.*

2. *Если линейное преобразование n-мерного пространства имеет n различных собственных значений, то его собственные векторы, соответствующие различным собственным значениям, образуют базис пространства.*

Доказательство. 1. ⊓ Существует $n+1$ различных собственных значений. Тогда им соответствует $n+1$ собственных векторов, которые по теореме линейно независимы. Но в n-мерном пространстве не существует $n+1$ линейно независимых векторов. ⊔

2. Здесь в n-мерном пространстве n собственных векторов линейно независимы и поэтому образуют базис. ■

Приведём утверждение, которое даёт возможность вычисления координат собственных векторов преобразования, если известны его собственные значения. Эта теорема не гарантирует наличие собственных значений. Как показывают примеры в начале этого раздела существуют преобразования, которые не имеют собственных значений.

Теорема 5.5.2. *Пусть в линейном пространстве E над полем P задан базис $e = (e_1, \ldots, e_n)$ и линейное преобразование A с матрицей* $[A]_{e,e} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$. *Преобразование A имеет собственный вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i$, отвечающий собственному значению λ_0, тогда и только тогда, когда*

• *λ_0 является корнем характеристического многочлена матрицы $[A]_{e,e}$*

$$f(\lambda) = \det([A]_{e,e} - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix}; \qquad (111)$$

• $[x]_e$ *— координатный столбец вектора* x *в базисе* e*, является нетривиальным решением матричного уравнения*

$$[A]_{e,e}[x]_e = \lambda_0 [x]_e \qquad (112)$$

или системы уравнений

$$\begin{cases} (a_{11}-\lambda_0)\beta_1 + a_{12}\beta_2 + \ldots + a_{1n}\beta_n = 0, \\ a_{21}\beta_1 + (a_{22}-\lambda_0)\beta_2 + \ldots + a_{2n}\beta_n = 0, \\ \qquad \ldots \\ a_{n1}\beta_1 + a_{n2}\beta_2 + \ldots + (a_{nn}-\lambda_0)\beta_n = 0. \end{cases} \qquad (113)$$

Доказательство. 1) Пусть преобразование A имеет собственный вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$, отвечающий собственному значению $\lambda = \lambda_0$. Тогда

$A(x) = \lambda_0 x \boxed{\text{приравняем координат-ные столбцы}} \Rightarrow [A(x)]_e = [\lambda_0 x]_e \boxed{\text{заменим } [A(x)]_e = [A]_{e,e}[x]_e \text{ и } [\lambda_0 x]_e = \lambda_0 [x]_e} \Rightarrow [A]_{e,e}[x]_e =$ $= \lambda_0 [x]_e \Rightarrow$ координатный столбец вектора $[x]_e$ является нетривиальным решением матричного уравнения $[A]_{e,e}[x]_e = \lambda_0 [x]_e$, т.е. выполнено (112). В подробной записи это равенство имеет вид

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix} = \lambda_0 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}.$$

Сделав умножение и приравняв соответствующие элементы левой и правой частей получим систему

$$\begin{cases} a_{11}\beta_1 + a_{12}\beta_2 + \ldots + a_{1n}\beta_n = \lambda_0\beta_1, \\ a_{21}\beta_1 + a_{22}\beta_2 + \ldots + a_{2n}\beta_n = \lambda_0\beta_2, \\ \dots \qquad \dots \qquad \dots \qquad \dots \qquad \dots \\ a_{n1}\beta_1 + a_{n2}\beta_2 + \ldots + a_{nn}\beta_n = \lambda_0\beta_n, \end{cases}$$

которая эквивалентна (113). Следовательно, столбец $(\beta_1, \beta_2, \ldots, \beta_n)^\top$ является нетривиальным решением системы (113).

2) Проведя рассуждения пункта 1) в обратном порядке, получим справедливость следующего утверждения: если столбец $(\beta_1, \ldots, \beta_n)^\top$ является нетривиальным решением системы (113), то вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq$ $\neq \theta$ удовлетворяет равенству $A(x) = \lambda_0 x$ и, следовательно, преобразование A имеет собственный вектор x отвечающий собственному значению $\lambda = \lambda_0$. ■

Теорема 5.5.3. а) *В n-мерном линейном пространстве E над полем P для произвольного линейного преобразования A характеристический многочлен* (111) *матрицы $[A]_{e,e} = (a_{ij})$ в базисе $e = (e_1, \ldots, e_n)$ является многочленом n-й степени от переменной $\lambda \in P$:*

$$f(\lambda) = (-\lambda)^n + \sum_{i=1}^{n} a_{ii}(-\lambda)^{n-1} + \sum_{i_1<i_2} \begin{vmatrix} a_{i_1,i_1} & a_{i_1,i_2} \\ a_{i_2,i_1} & a_{i_2,i_2} \end{vmatrix} (-\lambda)^{n-2} + \ldots,$$

причём $f(\lambda)$ не зависит от выбора базиса пространства.

б) *В n-мерном комплексном линейном пространстве E характеристический многочлен любого линейного преобразования A имеет ровно n корней, если каждый корень считать столько раз какова его кратность.*

в) *В n-мерном комплексном линейном пространстве E любое линейное преобразование имеет по крайней мере один собственный вектор.*

Доказательство. Пусть задан характеристический многочлен матрицы $[A]_{e,e} = (a_{ij})$ линейного преобразования A в базисе e :

$$f(\lambda) = \det([A]_{e,e} - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix}.$$

В этом определителе каждый диагональный элемент матрицы является многочленом от λ степени 1, каждый внедиагональные элемент не зависит от λ и имеет степень не выше 0. Отсюда следует, что диагональный член $d = (a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda)$ имеет степень, равную n, а каждый внедиагональный член определителя имеет меньшую степень. Поэтому $f(\lambda)$ является многочленом степени n и его старший коэффициент равен $(-1)^n$.

Найдём коэффициент многочлена $f(\lambda)$ при λ^{n-1}. Каждый внедиагональный член определителя содержит не более $n-2$ множителя из диагональных элементов, т. е. имеет степень, не большую $n-2$. Следовательно, в многочлене $f(\lambda)$ слагаемые, содержащие степень λ^{n-1}, определяются только диагональным членом $d=(a_{11}-\lambda)(a_{22}-\lambda)\ldots(a_{nn}-\lambda)$ и имеют вид $(a_{11}+a_{22}+\ldots+a_{nn})(-\lambda)^{n-1}$.

Аналогично найдём коэффициент многочлена $f(\lambda)$ при λ^{n-2} : возьмём $\sum\limits_{i_1<i_2} a_{i_1,i_1}a_{i_2,i_2}\dfrac{d}{(a_{i_1,i_1}-\lambda)(a_{i_2,i_2}-\lambda)}$ из диагонального члена d и $\sum\limits_{i_1<i_2}(-a_{i_1,i_2}a_{i_2,i_1})\dfrac{d}{(a_{i_1,i_1}-\lambda)(a_{i_2,i_2}-\lambda)}$ из внедиагональных членов, объединим эти суммы $\sum\limits_{i_1<i_2}\begin{vmatrix} a_{i_1,i_1} & a_{i_1,i_2} \\ a_{i_2,i_1} & a_{i_2,i_2}\end{vmatrix}\dfrac{(a_{11}-\lambda)(a_{22}-\lambda)\ldots(a_{nn}-\lambda)}{(a_{i_1,i_1}-\lambda)(a_{i_2,i_2}-\lambda)}$ и, выделив коэффициент при $(-\lambda)^{n-2}$, получим $\sum\limits_{i_1<i_2}\begin{vmatrix} a_{i_1,i_1} & a_{i_1,i_2} \\ a_{i_2,i_1} & a_{i_2,i_2}\end{vmatrix}(-\lambda)^{n-2}$.

Пусть для того же линейного преобразования A в другом базисе $e'=(e'_1,\ldots,e'_n)$ записан характеристический многочлен матрицы $[A]_{e',e'}$: $f_1(\lambda)=\det([A]_{e',e'}-\lambda I)$. Покажем, что $f_1(\lambda)=f(\lambda)$. Для этого воспользуемся из следствия теоремы 5.2.3 формулой (98) связи матриц линейного преобразования в двух базисах: $[A]_{e',e'}=T^{-1}_{e\to e'}[A]_{e,e}T_{e\to e'}$.

Имеем $f_1(\lambda)=\det([A]_{e',e'}-\lambda I)\boxed{\begin{array}{l}\text{заменим}\\ [A]_{e',e'}=T^{-1}_{e\to e'}[A]_{e,e}T_{e\to e'}\\ I=T^{-1}_{e\to e'}IT_{e\to e'}\end{array}}=$

$=\det(T^{-1}_{e\to e'}[A]_{e,e}T_{e\to e'}-\lambda T^{-1}_{e\to e'}IT_{e\to e'})\boxed{\begin{array}{l}\text{вынесем}\\ T^{-1}_{e\to e'}\ \text{влево и}\\ T_{e\to e'}\ \text{вправо}\end{array}}=$

$=\det(T^{-1}_{e\to e'}([A]_{e,e}-\lambda I)T_{e\to e'})\boxed{\begin{array}{l}\text{заменим}\ \det(CDE)=\\ =\det(C)\det(D)\det(E)\end{array}}=$

$=\det(T^{-1}_{e\to e'})\det([A]_{e,e}-\lambda I)\det(T_{e\to e'})\boxed{\begin{array}{l}\text{переставим множители}\\ \text{(значения определите-}\\ \text{лей)}\end{array}}=$

$=\det(T^{-1}_{e\to e'})\det(T_{e\to e'})\det([A]_{e,e}-\lambda I)\boxed{\begin{array}{l}\text{заменим}\ \det(C)\det(E)=\\ =\det(CE)\end{array}}$

$=\det(T^{-1}_{e\to e'}T_{e\to e'})\det([A]_{e,e}-\lambda I)=\det(I)\det([A]_{e,e}-\lambda I)=$

$= \det([A]_{e,e} - \lambda I) = f(\lambda)$.

б) В n-мерном комплексном линейном пространстве E характеристический многочлен произвольного линейного преобразования A является многочленом степени n над полем C и по следствию из основной теоремы алгебры многочленов имеет ровно n корней, если каждый корень считать столько раз какова его кратность.

в) Пусть λ_0 — корень характеристического многочлена (111), то есть $f(\lambda_0) = 0$. Рассмотрим систему (113). Матрицу этой системы можно представить в виде $B = D - \lambda_0 I$, где $D = [A]_{e,e}$ — матрица преобразования, I — единичная $n \times n$-матрица. Определитель матрицы B равен $f(\lambda_0) = 0$, поэтому система имеет ненулевое решение $(\beta_1, \beta_2, \ldots, \beta_n)^\top$.

Тогда по теореме 5.5.2 вектор $x = \sum\limits_{i=1}^{n} \beta_i e_i \neq \theta$ удовлетворяет равенству $A(x) = \lambda_0 x$ и, следовательно, преобразование A имеет собственный вектор x отвечающий собственному значению $\lambda = \lambda_0$. ■

Сформулируем *алгоритм поиска собственных значений и соответствующих собственных векторов.*

Требуется найти собственные векторы и собственные значения линейного преобразования A, т.е. $\lambda \in P$ и векторы $x \in E$ такие, что $A(x) = \lambda x,\ x \neq \theta$.

Начало.

1. Находим матрицу преобразования $[A]_{e,e} = (a_{ij})$ в базисе $e = (e_1, e_2, \ldots, e_n)$. (Желательно выбирать базис так, чтобы матрица была попроще.)

2. Составляем характеристический многочлен
$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix}$$

3. Ищем корни уравнение $f(\lambda) = 0$.

4. Если корней нет, то собственных значений и собственных векторов нет. **Конец.**

5. В противном случае находим все корни $\lambda = \lambda_j$, $j = \overline{1,k}$, $1 \leqslant k \leqslant n$.

6. Для $j = \overline{1,k}$ при $\lambda = \lambda_j$ записываем систему

$$\begin{cases} (a_{11}-\lambda)\beta_1 & + & a_{12}\beta_2 & +\ldots+ & a_{1n}\beta_n & = & 0, \\ a_{21}\beta_1 & + & (a_{22}-\lambda)\beta_2 & +\ldots+ & a_{2n}\beta_n & = & 0, \\ \ldots & & \ldots & \ldots & \ldots & & \ldots \\ a_{n1}\beta_1 & + & a_{n2}\beta_2 & +\ldots+ & (a_{nn}-\lambda)\beta_n & = & 0. \end{cases}$$

Находим общее решение $(\beta_1, \beta_2, \ldots, \beta_n) \neq (0, 0, \ldots, 0)$.

Записываем собственные векторы $x = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n$, соответствующие собственному значению $\lambda = \lambda_j$.

Конец. ■

5.6. Условия приведения матрицы линейного преобразования к диагональному виду. Жорданова форма матрицы линейного преобразования

Проанализируем возможность приведения матрицы линейного преобразования к диагональному виду, т.е. к наиболее простому виду.

Теорема 5.6.1. *Матрица линейного преобразования A линейного пространства E над полем P в некотором базисе $e = (e_1, e_2, \ldots, e_n)$ имеет диагональный вид*

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \tag{114}$$

тогда и только тогда, когда базис состоит из собственных векторов e_1, e_2, $\ldots$, e_n, соответствующих собственным знаменям λ_1, λ_2, $\ldots$, λ_n (собственные значения не обязательно различны).

Доказательство.

1) Пусть матрица преобразования A в базисе $e = (e_1, e_2, \ldots, e_n)$ имеет диагональный вид (114). По определению матрица состоит из координатных столбцов образов базисных векторов, т.е. из столбцов $[A(e_1)]_e$, $[A(e_2)]_e$, $\ldots$, $[A(e_n)]_e$. Следовательно, из столбцов матрицы (114) видим,

что i-я координата вектора $A(e_i)$ равна λ_i, а остальные координаты равны 0, тогда $A(e_i) = 0e_1 + \ldots + \lambda_i e_i + \ldots + 0e_n$, т.е. $A(e_i) = \lambda_i e_i$, $i = \overline{1,n}$. Итак, базис состоит из собственных векторов.

2) Пусть базис $e = (e_1, e_2, \ldots, e_n)$ состоит из собственных векторов и $A(e_i) = \lambda_i e_i$, $i = \overline{1,n}$. Тогда $A(e_i) = 0e_1 + \ldots + \lambda_i e_i + \ldots + 0e_n$. Записав координатные столбцы векторов $[A(e_i)]_e$ в матрицу $[A]_{e,e}$, видим,что она имеет диагональный вид (114) . ■

Определение 1. *Квадратная $k \times k$-матрица вида*

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

называется жордановой клеткой порядка k.

Клеточно-диагональная матрица, у которой все диагональные клетки жордановы, называется жордановой матрицей или матрицей, имеющей жорданову форму.

Замечания.

1. При $k = 1$ жордановой клеткой порядка 1 является матрица (λ).

2. Объединение всех диагональных элементов жордановой матрицы даёт все её собственные значения.

3. Диагональная матрица является частным случаем жордановой матрицы, имеющим все жордановы клетки порядка 1.

Теорема 5.6.2 (Жордан).

Для любого линейного преобразования конечномерного комплексного линейного пространства существует жорданов базис, в котором его матрица имеет жорданову форму. Жорданова форма матрицы единственна с точностью до порядка расположения диагональных клеток. ■

6. Евклидовы и унитарные пространства

6.1. Определение евклидова и унитарного пространства. Простейшие свойства скалярного произведения

Определение 1. *Линейное пространство E над полем $P = R$ или $P = C$ называется соответственно евклидовым или унитарным, если каждой паре векторов f, g из E поставлено в соответствие число из P, обозначаемое $(f,\, g)$ и называемое скалярным произведением* [7] *векторов f и g, причём выполнены следующие аксиомы:*

A1. $(f,\, g) = \overline{(g,\, f)}$, [8]

A2. $(f_1 + f_2,\, g) = (f_1,\, g) + (f_2,\, g)$,

A3. $(\alpha f,\, g) = \alpha(f,\, g)$,

A4. $(f,\, f) > 0$ *при* $f \neq \theta$,

(*здесь* f, f_1, f_2, g *— любые векторы пространства* E, α *— любое число из* P).

Замечание 1. Из аксиомы A1 при $g = f$ следует, что $(f,\, f) = \overline{(f,\, f)}$, т.е. $(f,\, f)$ — вещественное число, причём в силу A4 при $f \neq \theta$ — это положительное вещественное число.

Замечание 2. Для евклидова пространства скалярное произведение принимает только вещественные значения, поэтому аксиому A1 можно записать в виде: $(f,\, g) = (g,\, f)$.

Пример 1. В пространстве[9] $\mathbb{R}^n$ скалярное произведение векторов $x = (x_1,\, x_2,\, \ldots,\, x_n)$, $y = (y_1,\, y_2,\, \ldots,\, y_n)$ можно определить следующим образом:

$$(x, y) = \sum_{j=1}^{n} x_j y_j, \tag{115}$$

а в пространстве $\mathbb{C}^n$ —

$$(x, y) = \sum_{j=1}^{n} x_j \overline{y_j}. \tag{116}$$

[7] Само отображение, которое каждой паре векторов f, g из E ставит в соответствие их скалярное произведение, называют скалярным умножением. А скалярное произведение $(f,\, g)$ — это результат вычисления отображения на паре векторов f, g.

[8] Число $\overline{z} = x - yi$ называется сопряжённым для комплексного числа $z = x + yi$. Очевидно, что $\overline{x + 0i} = x - 0i$, т.е. сопряжённым для вещественного числа x является само это число.

[9] Пространство матриц-строк $\mathbb{R}^{1\times n}$ или матриц-столбцов $\mathbb{R}^{n\times 1}$.

Проверим выполнение аксиом для (116). Здесь $x,\, y,\, z \in E,\, \alpha \in P$.

1. $(y,\, x) = \sum\limits_{j=1}^{n} y_j \overline{x_j} = \overline{\sum\limits_{j=1}^{n} x_j \overline{y_j}} = \overline{(x,\, y)}$.

2. $(x+y,\, z) = \sum\limits_{j=1}^{n} (x_j + y_j)\overline{z_j} = \sum\limits_{j=1}^{n} x_j \overline{z_j} + \sum\limits_{j=1}^{n} y_j \overline{z_j} = (x,\, z) + (y,\, z)$.

3. $(\alpha x,\, y) = \sum\limits_{j=1}^{n} (\alpha x_j)\overline{y_j} = \alpha \sum\limits_{j=1}^{n} x_j \overline{y_j} = \alpha (x,\, y)$.

4. $(x,\, x) = \sum\limits_{j=1}^{n} x_j \overline{x_j} = \sum\limits_{j=1}^{n} |x_j|^2 > 0$ при $x \neq \theta$. ■

Пример 2. Рассмотрим функциональное пространство $C[0,1]$ (множество функций $f(x)$ таких, что $f\colon [0,1] \to \mathbb{R}$ и $f(x)$ непрерывна на отрезке $[0,1]$) с операциями сложения функций и умножения функций на вещественные числа. Здесь скалярное произведение функций $f = f(x)$ и $g = g(x)$ может быть задано так:

$$(f, g) = \int\limits_0^1 f(x) g(x)\, dx. \tag{117}$$

Легко проверяется, что формула (117) определяет скалярное произведение и выполнены аксиомы евклидова пространства. ■

Пример 3. Рассмотрим функциональное пространство $C([0,1], \mathbb{C})$. Это множество комплекснозначных функций $f(x) = u(x) + iv(x)$, $(u, v \in C[0,1])$ с операциями сложения функций и умножения функций на комплексные числа. Здесь скалярное произведение функций $f(x) = u(x) + iv(x)$ и $g(x) = p(x) + ir(x)$ может быть задано так:

$$(f, g) = \int\limits_0^1 f(x) \overline{g(x)}\, dx. \tag{118}$$

или после преобразования:

$$(f, g) = \int\limits_0^1 (u(x) + iv(x))(p(x) - is(x)\, dx = \int\limits_0^1 ((p(x)u(x) + s(x)v(x)) + i(p(x)v(x) - s(x)u(x)))\, dx = \int\limits_0^1 (p(x)u(x) + s(x)v(x))\, dx + i \int\limits_0^1 (p(x)v(x) - s(x)u(x))\, dx.$$

Легко проверяется, что формула (118) определяет скалярное произведение и выполнены аксиомы унитарного пространства. ■

Теорема 6.1.1 (Простейшие свойства скалярного умножения).

1. $\forall\, f,\, g_1,\, g_2 \in E \quad (f,\, g_1 + g_2) = (f,\, g_1) + (f,\, g_2)$.

2. $\forall\, f, g \in E,\ \beta \in P \quad (f,\, \beta g) = \overline{\beta}(f,\, g)$.

3. Формула скалярного произведения двух линейных комбинаций.

$\forall\, m, r \in \mathbb{N},\ f_1,\, \ldots,\, f_m,\, g_1,\, \ldots,\, g_r, \in E,\ \alpha_1,\, \ldots,\, \alpha_m,\, \beta_1,\, \ldots,\, \beta_r \in P$

$$\left(\sum_{i=1}^{m} \alpha_i f_i, \sum_{j=1}^{r} \beta_j g_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{r} \alpha_i \overline{\beta_j}(f_i, g_j). \tag{119}$$

4. $(\theta,\, g) = (g,\, \theta) = 0$ *для всех* $g \in E$ *и, в частности,* $(\theta,\, \theta) = 0$.

5. а) $(f,\, f) \geqslant 0$ *для всех* $f \in E$, б) $(f,\, f) = 0$ *тогда и только тогда, когда* $f = \theta$.

6. $(f,\, g) = 0$ *для всех* $g \in E$ тогда и только тогда, когда $f = \theta$.

7. *Любое подпространство евклидова (унитарного) пространства является евклидовым (унитарным) пространством.*

Доказательство.

1. $(f,\, g_1 + g_2) = \overline{(g_1 + g_2,\, f)} = \overline{(g_1,\, f) + (g_2,\, f)} = \overline{(g_1,\, f)} + \overline{(g_2,\, f)} = (f,\, g_1) + (f,\, g_2)$. Здесь использована аксиома А1 три раза, аксиома А2 один раз, и один раз утверждение: сопряжённое значение от суммы чисел равно сумме сопряжённых чисел.

2. $(f,\, \beta g) = \overline{(\beta g,\, f)} = \overline{\beta(g,\, f)} = \overline{\beta}\,\overline{(g,\, f)} = \overline{\beta}(f,\, g)$. Здесь использована аксиома А1 два раза, аксиома А3 один раз, и один раз утверждение: сопряжённое значение от произведения чисел равно произведению сопряжённых чисел.

3. Доказывается методом математической индукции.

4. Из аксиомы А3 при $\alpha = 0$ получаем $(0f,\, g) = 0(f,\, g)$. В силу $0f = \theta$ и $0(f,\, g) = 0$ выполнено равенство $(\theta,\, g) = 0$. Тогда по аксиоме А1 $(g,\, \theta) = \overline{(\theta,\, g)} = \overline{0} = 0$.

5. Выполнены условия $(f,\, f) > 0$ при $f \neq \theta$ по аксиоме А4 и $(\theta,\, \theta) = 0$ по утверждению 4. Тогда выполнены утверждение 5а: $(f,\, f) \geqslant 0$ при всех $f \in E$ и утверждение 5б: из $f = \theta$ следует, что $(f,\, f) = 0$, а из условия $(f,\, f) = 0$ следует, что обязательно $f = \theta$.

6. Если $(f,\, g) = 0$ при всех $g \in E$, то для $g = f$ получаем $(f,\, f) = 0$ и из условия 5б) $f = \theta$. Если $f = \theta$, то из утверждения 4 следует $(f,\, g) = 0$ при всех $g \in E$.

7. Скалярное произведение определено на каждой паре векторов пространства и тогда на каждой паре векторов подпространства. Линейное подпространство замкнуто относительно операций сложения и умножения на числа, поэтом аксиомы A1–A4 выполнены на всех векторах подпространства и для всех чисел из соответствующего поля. ■

6.2. Неравенство Коши–Буняковского. Длина вектора

В этом разделе рассматривается евклидово (унитарное) пространство E над полем $P = \mathbb{R}$ ($P = \mathbb{C}$).

Теорема 6.2.1 (Неравенство Коши-Буняковского).

1° *Для всех векторов* $f,\, g \in E$

$$|(f,\, g)|^2 \leqslant (f,\, f)(g,\, g). \tag{120}$$

или, в эквивалентной форме

$$\begin{vmatrix} (f,\, f) & (f,\, g) \\ (g,\, f) & (g,\, g) \end{vmatrix} \geqslant 0.$$

2° *Неравенство (120) обращается в равенство тогда и только тогда, когда векторы* $f,\, g$ *линейно зависимы (пропорциональны).*

Доказательство. 1° Берём любые векторы $f,\, g \in E$.

При $g = \theta$ имеем $(f,\, g) = 0$, $(g, g) = 0$ и неравенство (120) обращается в равенство.

При $g \neq \theta$ и произвольном $\beta \in P$ рассмотрим скалярное произведение $(f - \beta g, f - \beta g)$. В силу утверждения 5а теоремы 6.1.1 это скалярное произведение принимает неотрицательное значение, т.е.

$$(f - \beta g, f - \beta g) \geqslant 0.$$

Преобразуя скалярное произведение по формуле (119), получим неравенство

$$(f,\, f) - \overline{\beta}(f,\, g) - \beta(g,\, f) + \beta\overline{\beta}(g,\, g) \geqslant 0.$$

При $\beta = \dfrac{(f,\, g)}{(g,\, g)}$ в силу $(g,\, g) > 0$ имеем $\overline{\beta} = \dfrac{\overline{(f,\, g)}}{\overline{(g,\, g)}} = \dfrac{\overline{(f,\, g)}}{(g,\, g)}$, тогда

$\overline{\beta}(f,\, g) = \dfrac{\overline{(f,\, g)}(f,\, g)}{(g,\, g)} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$, $\beta(g,\, f) = \beta\overline{(f,\, g)} = \dfrac{(f,\, g)\overline{(f,\, g)}}{(g,\, g)} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$,

$\beta\overline{\beta}(g,\, g) = \dfrac{(f,\, g)\overline{(f,\, g)}(g,\, g)}{(g,\, g)^2} = \dfrac{|(f,\, g)|^2}{(g,\, g)}$, и неравенство примет вид:

$$(f,\, f) - \frac{|(f,\, g)|^2}{(g,\, g)} - \frac{|(f,\, g)|^2}{(g,\, g)} + \frac{|(f,\, g)|^2}{(g,\, g)} \geqslant 0 \quad \text{или}$$

$$(f,\, f) - \frac{|(f,\, g)|^2}{(g,\, g)} \geqslant 0.$$

Последнее неравенство запишем в виде

$$\frac{|(f,\, g)|^2}{(g,\, g)} \leqslant (f,\, f),$$

откуда после умножения на $(g, g) > 0$ получается (120).

2° Часть утверждения фактически получается из рассуждений при доказательстве неравенства (120).

Пусть неравенство (120) обратилось в равенство.

При $g = \theta$ векторы f, g линейно зависимы и $g = 0f$.

При $g \neq \theta$ выполнено $(g, g) \neq 0$. Тогда проходим неравенства в рассуждениях при доказательстве 1° в обратном порядке, начиная с равенства (120), и все они обращаются в равенства. В итоге получаем равенство $(f - \beta g, f - \beta g) = 0$ при $\beta = \dfrac{(f,\, g)}{(g,\, g)}$. Отсюда следует, что $f - \beta g = \theta$, т.е. векторы f, g линейно зависимы и $f = \beta g$.

Пусть $f,\, g$ линейно зависимы.

При $g = \theta$ или $g = \theta$ неравенство (120) обращается в равенство.

При $f \neq \theta$ и $g \neq \theta$ существует нетривиальная линейная комбинация, равная θ : $\gamma f + \delta g = \theta$. Здесь $\gamma \neq 0$. В противном случае $\gamma = 0$, $\delta \neq 0$ и равенство $0f + \delta g = \theta$ противоречит $g \neq \theta$. Аналогично, $\delta \neq 0$. Далее из равенства $\gamma f + \delta g = \theta$ получим $f + \dfrac{\delta}{\gamma} g = \theta$ и, обозначив $\beta = -\dfrac{\delta}{\gamma}$, запишем $f - \beta g = \theta$ и $f = \beta g$.

Найдём значение β. Умножив равенство $f = \beta g$ скалярно на g, получим $(f, g) = (\beta g, g)$ или $(f, g) = \beta(g, g)$. Отсюда выражаем $\beta = \dfrac{(f, g)}{(g, g)}$.

Итак, при $\beta = \dfrac{(f, g)}{(g, g)}$ имеем $f - \beta g = \theta$ и $(f - \beta g, f - \beta g) = 0$, то есть неравенство $(f - \beta g, f - \beta g) \geqslant 0$ обратилось в равенство. Тогда все неравенства в рассуждениях при доказательстве утверждения 1°, включая само неравенство (120), обращаются в равенства. ■

Определение 1. *Длиной вектора f называется арифметическое значение квадратного корня из его скалярного квадрата (f, f).*

Обозначение: $\|f\|$ *или* $|f|$.

Итак, $\|f\| = \sqrt{(f, f)}$.

Длину вектора также называют нормой.

Вектор единичной длины называется нормированным.

Теорема 6.2.2 (Свойства длины).

1° *Для всех векторов* $f \in E$ *существует длина* $\|f\| = \sqrt{(f,f)}$.

2° $\forall f \in E \quad \|f\| = \sqrt{(f,f)} \Leftrightarrow \|f\|^2 = (f,f)$.

3° *Неравенство Коши-Буняковского (120) можно записать в виде*

$$|(f,g)| \leqslant \|f\|\|g\|.$$

4° $\forall f \in E \quad \|f\| \geqslant 0$ *и* $\|f\| = 0 \Leftrightarrow f = \theta$.

5° $\|\alpha f\| = |\alpha|\|f\|$ *для всех векторов* $f \in E$ *и чисел* $\alpha \in P$.

6° $\|f+g\| \leqslant \|f\| + \|g\|$ *для всех векторов* $f, g \in E$.

Причём знак равенства здесь достается тогда и только тогда, когда $f = \beta g$ *или* $g = \beta f$ *при некотором* $\beta \geqslant 0$.

7° $\big|\|f\| - \|g\|\big| \leqslant \|f+g\|$ *для всех векторов* $f, g \in E$.

Причём знак равенства здесь достается тогда и только тогда, когда $f = \beta g$ *или* $g = \beta f$ *при некотором* $\beta \leqslant 0$.

Доказательство. 1°, 2°, 4° Для любого вектора $f \in E$ в силу утверждения 5 теоремы 6.1.1 $(f,f) \geqslant 0$, причём $(f,f) = 0 \Leftrightarrow f = \theta$. Следовательно, а) существует длина $\|f\| = \sqrt{(f,f)}$, б) $\|f\|^2 = (f,f)$, в) $\|f\| \geqslant 0$ и $\|f\| = 0 \Leftrightarrow f = \theta$.

3° $|(f,g)|^2 \leqslant (f,f)(g,g) \Leftrightarrow |(f,g)|^2 \leqslant \|f\|^2\|g\|^2 \Leftrightarrow |(f,g)| \leqslant \|f\|\|g\|$. ∎

5° $\|\alpha f\|^2 = (\alpha f, \alpha f) = \alpha\overline{\alpha}(f,f) = |\alpha|^2\|f\|^2 \Leftrightarrow \|\alpha f\|^2 = |\alpha|^2\|f\|^2 \Leftrightarrow \|\alpha f\| = |\alpha|\|f\|$.

6° Покажем, что $\|f+g\|^2 \leqslant (\|f\|+\|g\|)^2$, тогда $\|f+g\| \leqslant \|f\|+\|g\|$. Из $\|f+g\|^2 = (f+g, f+g)$ получим $\|f+g\|^2 = (f,f)+(f,g)+(g,f)+(g,g)$. Возьмём модуль левой и правой частей последнего равенства, учитывая, что они неотрицательны, и воспользуемся неравенством модуль суммы не превосходит суммы модулей, тогда $\|f+g\|^2 \leqslant |(f,f)|+|(f,g)|+|(g,f)|+|(g,g)|$. Используя 2° и 3°, придём к неравенству $\|f+g\|^2 \leqslant \|f\|^2 + \|f\|\|g\| + \|g\|\|f\| + \|g\|^2$. Откуда $\|f+g\|^2 \leqslant \|f\|^2 + 2\|f\|\|g\| + \|g\|^2$ или $\|f+g\|^2 \leqslant (\|f\|+\|g\|)^2$.

7° Покажем, что $\|f+g\|^2 \geqslant (\|f\|-\|g\|)^2$, тогда $\|f+g\| \geqslant \big|\|f\| - \|g\|\big|$.

Из доказательства 6° возьмём $\|f+g\|^2 = |(f,f)+(g,g)+(f,g)+(g,f)|$, воспользуемся два раза неравенством $|z_1+z_2| \geqslant |z_1|-|z_2|$ и получим $\|f+g\|^2 \geqslant (f,f)+(g,g)-|(f,g)|-|(g,f)|$. Оценим снизу правую часть последнего неравенства. Из неравенства Коши-Буняковского $|(f,g)| \leqslant \|f\|\|g\|$ и $|(g,f))| \leqslant \|f\|\|g\|$, тогда $-|(f,g)|-|(g,f)| \geqslant -2\|f\|\|g\|$. Следовательно, получаем неравенство $\|f+g\|^2$
$geqslant \|f\|^2 - 2\|f\|\|g\| + \|g\|^2$ или $\|f+g\|^2 \geqslant (\|f\|-\|g\|)^2$. ■

Замечание. Любой ненулевой вектор f можно нормировать $\dfrac{f}{\|f\|}$, разделив его на длину $\|f\|$, или $\dfrac{1}{\|f\|}f$, умножив на $\dfrac{1}{\|f\|}$.

Определение 2. *Неравенства*

$$\big|\|f\|-\|g\|\big| \leqslant \|f+g\| \ \text{ и } \ \|f+g\| \leqslant \|f\|+\|g\|$$

называют неравенствами треугольника.[10]

Определение 3. *Углом между между ненулевыми векторами f и g в евклидовом пространстве назовём тот угол φ $(0 \leqslant \varphi \leqslant \pi)$, косинус которого равен отношению $\dfrac{(f,g)}{\|f\|\cdot\|g\|}$, т.е. $\cos\varphi = \dfrac{(f,g)}{\|f\|\cdot\|g\|}$.*

Это определение угла корректно, так как скалярное произведение в евклидовом пространстве принимает только вещественные значения и из неравенства Коши-Буняковского следует, что $-1 \leqslant \dfrac{(f,g)}{\|f\|\cdot\|g\|} \leqslant +1$.

В унитарном пространстве скалярное произведение принимает комплексные значения, поэтому понятия угла там не вводится. Рассматривается лишь понятие ортогональности.

Задача 6.2.1. *Напишите как выглядит неравенство Коши-Буняковского в пространствах примеров 1-3 на стр. 180–181*

6.3. Ортогональность

В этом разделе рассмотрим понятие ортогональности в евклидовом (унитарном) пространстве E.

[10] В любом треугольнике ABC выполняются неравенства $\big||\overline{AB}|-|\overline{BC}|\big| \leqslant |\overline{AC}|$ и $|\overline{AC}| \leqslant |\overline{AB}|+|\overline{BC}|$.

Определение 1. *Два вектора f и g пространства E называются ортогональными, если $(f, g) = 0$. Обозначение: $f \perp g$.*

Система векторов $e_1, e_2, \ldots, e_k \in E$ называется ортогональной системой, если

$$(e_i, e_j) = 0 \quad \text{при } i \neq j. \tag{121}$$

Система векторов $e_1, e_2, \ldots, e_k \in E$ называется ортонормированной системой, если она является ортогональной системой и каждый вектор системы — нормированный, т.е.

$$(e_i, e_j) = \delta_{ij}, \tag{122}$$

где δ_{ij} — символ Кронекера, для которого $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$ ■

Теорема 6.3.1.

Ортогональная система ненулевых векторов линейно независима.

Доказательство. Пусть $e_1, e_2, \ldots, e_k \in E$ — ортогональная система ненулевых векторов и линейная комбинация векторов этой системы равна θ : $\sum\limits_{i=1}^{k} \alpha_i e_i = \theta$. Умножим обе части этого равенства скалярно на e_j : $\left(\sum\limits_{i=1}^{k} \alpha_i e_i, e_j\right) = (\theta, e_j)$. С помощью формулы скалярного произведения двух линейных комбинаций (119) находим $\sum\limits_{i=1}^{k} \alpha_i (e_i, e_j) = 0$. Здесь в силу ортогональности системы имеем $(e_i, e_j) = 0$ при $i \neq j$. Поэтому в сумме остаётся только одно слагаемое при $i = j$: $\alpha_j (e_j, e_j) = 0$. Поскольку $e_j \neq \theta$, то $(e_j, e_j) \neq 0$ и тогда $\alpha_j = 0$. Следовательно, все коэффициенты линейной комбинации равны нулю и векторы этой системы линейно независимы.

Следствия. 1. *Ортонормированная система векторов линейно независима.*

2. *В n-мерном пространстве E любая ортогональная система из n ненулевых векторов образует базис.* ■

Определение 2.

Базис, векторы которого образуют ортогональную (ортонормированную) систему называется ортогональным (ортонормированным) базисом. ■

Теорема 6.3.2. *Координаты* $\alpha_1, \ldots, \alpha_n$ *вектора* $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ *в базисе* $e = (e_1, \ldots, e_n)$ *вычисляются по правилу*

$$1) \quad \alpha_i = \frac{(f, e_i)}{(e_i, e_i)}, \quad i = \overline{1, n} \tag{123}$$

тогда и только тогда, когда e *— ортогональный базис;*

$$2) \quad \alpha_i = (f, e_i), \quad i = \overline{1, n} \tag{124}$$

тогда и только тогда, когда e *— ортонормированный базис.*

Доказательство. I. а) Пусть координаты вычисляются по правилу (123) для любого вектора $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и, в частности, для каждого $f = e_j$. Разложение вектора e_j по базису e имеет вид $e_j = e_j$, т. е. $e_j = \sum\limits_{i=1}^{n} \alpha_i e_i$, где $\alpha_i = 0$ при $i \neq j$ и $\alpha_j = 1$. Тогда из (123) получаем $0 = \dfrac{(e_j, e_i)}{(e_i, e_i)}$ при $i \neq j$, т.е. $(e_j, e_i) = 0$ при всех $i \neq j$ и e — ортогональный базис.

б) Докажем обратное утверждение. Пусть $e = (e_1, \ldots, e_n)$ — ортогональный базис и $f = \sum\limits_{i=1}^{n} \alpha_i e_i$. Возьмём произвольное $j \in \{1, \ldots, n\}$ и найдём скалярное произведение $(f, e_j) = \left(\sum\limits_{i=1}^{n} \alpha_i e_i, e_j\right) = \sum\limits_{i=1}^{n} \alpha_i (e_i, e_j)$. В силу ортогональности системы $(e_i, e_j) = 0$ при $i \neq j$. Следовательно, от суммы остаётся только j-е слагаемое, и мы получаем $(f, e_j) = \alpha_j (e_j, e_j)$ или $\alpha_j = \dfrac{(f, e_j)}{(e_j, e_j)}$ поскольку $(e_j, e_j) \neq 0$. Итак, координаты вычисляются по правилу (123).

II. а) Пусть координаты вычисляются по правилу (124) для любого вектора $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и, в частности, для каждого $f = e_j$. Для e_j имеем $e_j = \sum\limits_{i=1}^{n} \alpha_i e_i$, где $\alpha_i = 0$ при $i \neq j$ и $\alpha_j = 1$. Тогда из (124) получаем

$0 = (e_j, e_i)$ при $i \neq j$ и $1 = (e_j, e_j)$, т. е. e — ортонормированный базис.

б) Докажем обратное утверждение. Пусть $e = (e_1, \ldots, e_n)$ — ортонормированный базис и $f = \sum\limits_{i=1}^{n} \alpha_i e_i$. Тогда e — ортогональный базис и по правилу (123) имеем $\alpha_i = \dfrac{(f, e_i)}{(e_i, e_i)}$, а поскольку $(e_i, e_i) = 1$ выполнено правило (124).

Следствие. *Произвольный вектор f в базисе $e = (e_1, e_2, \ldots, e_n)$ имеет разложение*

$$f = \sum_{i=1}^{n} \frac{(f, e_i)}{(e_i, e_i)} e_i \quad \textit{для ортогональном базиса } e \textit{ и}$$

$$f = \sum_{i=1}^{n} (f, e_i)\, e_i \quad \textit{для ортонормированного базиса } e.$$

Доказательство. В разложение $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ подставим координаты α_i из формул (123) и (124). ■

Теорема 6.3.3.

Скалярное произведение векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $g = \sum\limits_{i=1}^{n} \beta_i e_i$, заданных своими координатами $\alpha_1, \ldots, \alpha_n$ и $\beta_1, \ldots, \beta_n$ в базисе $e = (e_1, \ldots, e_n)$, вычисляется по правилу

$$(f, g) = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}, \tag{125}$$

тогда и только тогда, когда e — ортонормированный базис.

Доказательство. а) Пусть скалярное произведение любых векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $g = \sum\limits_{i=1}^{n} \beta_i e_i$ вычисляется по правилу (125). Применив это правило для базисных векторов $e_j = \sum\limits_{i=1}^{n} \delta_{ij} e_i$ и $e_k = \sum\limits_{i=1}^{n} \delta_{ik} e_i$ получим $(e_j, e_k) = \sum\limits_{i=1}^{n} \delta_{ij} \overline{\delta_{ik}} = \sum\limits_{i=1}^{n} \delta_{ij} \delta_{ik} = \delta_{jj} \delta_{jk} = \delta_{jk}$, т.е. $e = (e_1, e_2, \ldots, e_n)$ — ортонормированный базис.

б) Докажем обратное утверждение.

Пусть $e = (e_1, \ldots, e_n)$ — ортонормированный базис. Используя формулу (119) скалярного произведения двух линейных комбинаций, вы-

числим скалярное произведение векторов $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ и $g = \sum\limits_{i=1}^{n} \beta_i e_i$: $\left(\sum\limits_{i=1}^{n} \alpha_i e_i, \sum\limits_{j=1}^{n} \beta_j e_j\right) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \alpha_i \overline{\beta_j}(e_i, e_j) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \alpha_i \overline{\beta_j}\delta_{ij} = \sum\limits_{i=1}^{n} \alpha_i \overline{\beta_i}$. Следовательно, скалярное произведение любых векторов вычисляется по правилу (125) . ■

Замечание. Для евклидова пространства в силу $\beta_i = \overline{\beta_i}$ формулу (125) можно записать в виде

$$(f, g) = \sum_{i=1}^{n} \alpha_i \beta_i. \ ■$$

6.4. Метод ортогонализации Грама–Шмидта. Построение ортогонального (ортонормированного) базиса

В этом разделе рассмотрим построение ортогональной системы векторов в евклидовом (унитарном) пространстве E и построение ортогонального (ортонормированного) базиса.

Метод ортогонализации.

Это некоторый способ перехода от заданной линейно независимой системы из k векторов $e = (e_1, \ldots, e_k)$ к ортогональной системе из k ненулевых векторов $f = (f_1, \ldots, f_k)$.

Применим метод математической индукции по $p \in N$ для построения из векторов системы $e_1, \ldots, e_p$ ортогональной системы из ненулевых векторов $f_1, \ldots, f_p$.

При $p = 1$ положим $f_1 = e_1$. Так как e_1 линейно независим, то вектор $f_1 \neq \theta$ и f_1 является линейной комбинацией вектора e_1.

При $p = 2$ возьмём $f_1 = e_1$ и положим $f_2 = \alpha_1 f_1 + e_2$. Так как $f_1 = e_1$, а векторы e_1, e_2 линейно независимы, то вектор $f_2 \neq \theta$ при любом числе α_1. Число α_1 найдём из условия ортогональности f_2 и f_1 :

$$0 = (f_2, f_1) = (\alpha_1 f_1 + e_2, f_1) = \alpha_1 (f_1, f_1) + (e_2, f_1).$$

В силу $(f_1, f_1) \neq 0$ получим $\alpha_1 = -\dfrac{(e_2, f_1)}{(f_1, f_1)}$. Тогда f_1 и f_2 ненулевые ортогональные векторы. Причём вектор f_1 есть линейная комбинация вектора e_1, вектор f_2 является линейной комбинацией векторов e_1, e_2.

Пусть при $p = m\,(m < k)$ уже построена ортогональная система ненулевых векторов $f_1, \ldots, f_m$, причём для каждого $i = \overline{1,m}$ вектор f_i является линейной комбинацией векторов $e_1, \ldots, e_i$.

При $p = m+1$ для первых m векторов из $e_1, \ldots, e_m, e_{m+1}$ по предположению индукции существует ортогональная система ненулевых векторов $f_1, \ldots, f_m$ таких, что для каждого $i = \overline{1,m}$ вектор f_i является линейной комбинацией векторов $e_1, \ldots, e_i$. Построим вектор f_{m+1} в виде

$$f_{m+1} = \alpha_1 f_1 + \ldots + \alpha_m f_m + e_{m+1}, \text{ где } \alpha_i \text{ — числа.}$$

Здесь для каждого $i = \overline{1,m+1}$ вектор f_i является линейной комбинацией векторов $e_1, \ldots, e_i$. Вектор $f_{m+1} = \sum\limits_{i=1}^{m} \alpha_i f_i + e_{m+1}$ не равен θ при любых числах α_i $(i = \overline{1,m})$, так как векторы $\sum\limits_{i=1}^{m} \alpha_i f_i$ линейно выражаются через $e_1, \ldots, e_m$, а вектор e_{m+1} линейно не выражается через $e_1, \ldots, e_m$ в силу линейной независимости векторов $e_1, \ldots, e_{m+1}$. Числа α_i $(i = \overline{1,m})$ найдём из условия ортогональности вектора f_{m+1} всем векторам f_i $(i = \overline{1,m})$:

$$0 = (f_{m+1}, f_i) = (\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_m f_m + e_{m+1}, f_i) =$$
$$= \alpha_1 (f_1, f_i) + \alpha_2 (f_2, f_i) + \ldots + \alpha_m (f_m, f_i) + (e_{m+1}, f_i).$$

Отсюда в силу ортогональности векторов $f_1, \ldots, f_m$ имеем $(f_j, f_i) = \delta_{ji}$.

Тогда $\alpha_i (f_i, f_i) + (e_{m+1}, f_i) = 0 \;\Rightarrow\; \alpha_i = -\dfrac{(e_{m+1}, f_i)}{(f_i, f_i)}$, $i = \overline{1,m}$ и

$$f_{m+1} = \sum_{i=1}^{m} \alpha_i f_i + e_{m+1}.$$

Таким образом, за k шагов мы построим ортогональную систему ненулевых векторов $f = (f_1, \ldots, f_k)$. ■

Алгоритм метод ортогонализации.

Начало. Полагаем $f_1 = e_1$.

Для $p = \overline{1, k-1}$ полагаем

$$f_{p+1} = \alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_p f_p + e_{p+1}, \text{ где } \alpha_i = -\frac{(e_{p+1}, f_i)}{(f_i, f_i)}, \quad i = \overline{1,p}.$$

Конец. ■

Для построения ортонормированной системы $g_1, g_2, \ldots, g_k$ можно построить сначала ортогональную систему $f_1, f_2, \ldots, f_k$, затем нормировать каждый вектор $g_i = \dfrac{f_i}{\|f_i\|}$, $i = \overline{1,k}$. ■

Теорема 6.4.1.

1° *В ненулевом конечномерном пространстве E существует ортогональный базис.*

2° *В ненулевом конечномерном пространстве E существует ортонормированный базис.*

Доказательство. Пусть $e = (e_1, \ldots, e_n)$ базис в E. Применим к системе e линейно независимых векторов метод ортогонализации и получим ортогональную систему, состоящую из n ненулевых векторов $f_1, \ldots, f_n$. В силу теоремы 6.3.1 эти n векторов линейно независимы и, следовательно, образуют ортогональный базис. Нормируем векторы $g_i = \dfrac{f_i}{\|f_i\|}$, $i = \overline{1,n}$ и получим ортонормированный базис $g_1, \ldots, g_n$. ■

6.5. Ортогональное дополнение. Разложение пространства на сумму ортогональных подпространств. Задача о перпендикуляре

Определение 1. *Подмножества U, V унитарного (евклидова) пространства E называются ортогональными, если любые векторы $x \in U$, $y \in V$ ортогональны, т. е. $(x, y) = 0$.*

Обозначение: $U \perp V$. Для одноточечного множества $U = \{x\}$ пишем $x \perp V$.

Множество всех векторов из E ортогональных U называется ортогональным дополнением множества U и обозначается $U^{\perp}$.

Лемма 6.5.1. 1° *Ортогональное дополнение любого непустого множества унитарного пространства E является подпространством.*

2° *Вектор x ортогонален векторам $u_1, u_2, \ldots, u_m$ $\Leftrightarrow$ x ортогонален к любой линейной комбинации этих векторов $\sum\limits_{i=1}^{m} \alpha_i u_i$.*

3° *Вектор x ортогонален векторам $u_1, u_2, \ldots, u_m$ $\Leftrightarrow$ x ортогонален линейной оболочке* $\mathrm{Lin}(u_1, u_2, \ldots, u_m)$.

4° *Вектор x ортогонален векторам базиса подпространства $L \Leftrightarrow$ x ортогонален L.*

5° *Два подпространства с базисами ортогональны тогда и только тогда, когда их базисы ортогональны.*

Доказательство.

1° Пусть $U \subseteq E$ и $U \neq \emptyset$. Применим к ортогональному дополнению $U^{\perp}$ критерий подпространства.

а) $x_1, x_2 \in U^{\perp} \Rightarrow x_1 \perp U, x_2 \perp U \Rightarrow \forall y \in U : (x_1, y) = 0, (x_2, y) = 0 \Rightarrow \forall y \in U : (x_1 + x_2, y) = 0 \Rightarrow x_1 + x_2 \perp U \Rightarrow x_1 + x_2 \in U^{\perp}$.

б) $x \in U^{\perp}, \alpha$ — число $\Rightarrow x \perp U \Rightarrow \forall y \in U : (x, y) = 0 \Rightarrow \forall y \in U : (\alpha x, y) = 0 \Rightarrow \alpha x \perp U \Rightarrow \alpha x \in U^{\perp}$.

2° $x \perp \{u_1, u_2, \ldots, u_m\} \Rightarrow (x, u_i) = 0, i = \overline{1, m} \Rightarrow \forall$ чисел α_i $(x, \sum\limits_{i=1}^{m} \alpha_i u_i) = \sum\limits_{i=1}^{m} \overline{\alpha}_i (x, u_i) = \sum\limits_{i=1}^{m} \overline{\alpha}_i 0 = 0 \Rightarrow \forall$ чисел α_i $x \perp \sum\limits_{i=1}^{m} \alpha_i u_i$.

Обратно, $\forall$ чисел α_i $x \perp \sum\limits_{i=1}^{m} \alpha_i u_i \Rightarrow$ для каждого $j = \overline{1, m}$ берём числа $\alpha_i = \delta_{ij}$ и получаем $x \perp \sum\limits_{i=1}^{m} \alpha_i u_i = u_j$.

3° эквивалентно 2°, т. к. $\mathrm{Lin}(u_1, u_2, \ldots, u_m) = \{x \colon x = \sum\limits_{i=1}^{m} \alpha_i u_i\}$.

4° следует из 3°, т. к. подпространство L есть линейная оболочка его базиса.

5° Пусть U, V — подпространства с базисами $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ соответственно.

Если $U \perp V$, то любые векторы $x \in U, y \in V$ ортогональны, т. е. $x \perp y$. В частности, $u_i \perp v_j$ для всех i, j.

Обратно, $u_i \perp v_j$ для всех i, j. В силу 2° произвольный вектор x из U, $x = \sum\limits_{i=1}^{n} \alpha_i u_i$ ортогонален каждому вектору базиса v_j, $j = \overline{1, n}$. Тогда в силу 4° этот произвольный вектор $x \in U$ ортогонален подпространству V. Следовательно, $U \perp V$. ■

Теорема 6.5.1. (Разложение пространства на прямую сумму ортогональных подпространств.)

Произвольное n-мерное унитарное пространство E есть прямая сумма любого подпространства U и ортогонального дополнения $U^{\perp}$:

$$E = U \dot{+} U^{\perp}. \tag{126}$$

Доказательство. Обозначим $k = \dim(U)$. Если $k = 0$, то $U = \theta$ и $U^{\perp} = E$, а если $k = n$, то $U = E$ и $U^{\perp} = \theta$. В этих случаях (126) выполнено.

Пусть $U \neq \theta$ и $U \neq E$, тогда $0 < k < n$. Выберем базис $e_1, \ldots, e_k$ в U и дополним его до базиса $e = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ в E. Применим к базису e метод ортогонализации, нормируем полученные векторы и получим ортонормированный базис $f = (f_1, \ldots, f_k, f_{k+1}, \ldots, f_n)$ в E. Из алгоритма метода следует, что для каждого i вектор f_i есть линейная комбинация векторов $e_1, \ldots, e_i$. Тогда векторы $f_1, \ldots, f_k$ принадлежат подпространству U и образуют в нём ортонормированный базис, который обозначим f'. Обозначим $W = \mathrm{Lin}(f_{k+1}, \ldots, f_n)$, — это подпространство в E с ортонормированным базисом $f'' = (f_{k+1}, \ldots, f_n)$. Объединение базисов f' и f'' является базисом суммы $U+W$, поэтому эта сумма прямая. А поскольку $U \dot{+} W \subseteq E$ и $\dim(U \dot{+} W) = n = \dim(E)$, то $E = U \dot{+} W$. В силу леммы 6.5.1 $U \perp W$, так как базисы ортогональны: $f' \perp f''$.

Покажем, что $W = U^{\perp}$.

1) $W \subseteq U^{\perp}$. Действительно, $\begin{cases} x \in W, \\ U \perp W, \end{cases} \Rightarrow x \perp U \Rightarrow x \in U^{\perp}$.

2) $U^{\perp} \subseteq W$. Действительно, $x \in U^{\perp} \Rightarrow x \perp U \Rightarrow x \perp f' \Rightarrow (x, f_i) = 0$, $i = \overline{1,k} \Rightarrow x = \sum\limits_{i=1}^{n} (x, f_i) f_i = \sum\limits_{i=k+1}^{n} (x, f_i) f_i \in W$.

Отсюда $W = U^{\perp}$ и тогда $E = U \dot{+} W = U \dot{+} U^{\perp}$. ■

Следствие. *Если U — подпространство n-мерного унитарного пространства E, то для любого вектора $f \in E$ существует, и притом единственное разложение*

$$f = g + h, \quad g \in U, \ h \perp U. \tag{127}$$

Доказательство следует из разложения (126) $E = U \dot{+} U^{\perp}$.

Определение 2. *Вектор g в разложении* (127) *называется ортогональной проекцией вектора f на подпространство U, а вектор h — ортогональной составляющей вектора f.*

Задачу разложения (127) *вектора на ортогональную проекцию и ортогональную составляющую называют задачей о перпендикуляре.*

Замечание. Часто вектор h называют *перпендикуляром, опущенным из конца вектора f на подпространство U.*

Решение задачи о перпендикуляре.

Из теоремы 6.5.1 следует, что задача о перпендикуляре имеет, и притом единственное, решение. Укажем один из способов построения этого решения.

Пусть заданы $a_1, \ldots, a_k \in E$ и $U = \mathrm{Lin}(a_1, \ldots, a_k)$. Для заданного вектора $f \in E$ ищем векторы 1) g и 2) h такие, что

$$\begin{cases} f = g + h, \\ g \in U, \\ h \perp U \end{cases} \Leftrightarrow \begin{cases} g = \sum\limits_{j=1}^{k} x_j a_j, \\ f - g \perp U, \\ h = f - g, \end{cases} \boxed{\begin{array}{l}\text{здесь } x_j, \\ j = \overline{1,k} \text{ —} \\ \text{числа}\end{array}} \Leftrightarrow \begin{cases} g = \sum\limits_{j=1}^{k} x_j a_j, \\ f - g \perp a_i,\ i = \overline{1,k}, \\ h = f - g \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} (f - g, a_i) = 0,\ i = \overline{1,k}, \\ g = \sum\limits_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases} \Leftrightarrow \begin{cases} (g, a_i) = (f, a_i),\ i = \overline{1,k}, \\ g = \sum\limits_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \Big(\sum\limits_{j=1}^{k} x_j a_j, a_i\Big) = (f, a_i),\ i = \overline{1,k}, \\ g = \sum\limits_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases} \Leftrightarrow \begin{cases} \sum\limits_{j=1}^{k} x_j (a_j, a_i) = (f, a_i), \\ \qquad i = \overline{1,k}, \\ g = \sum\limits_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases}$$

Итак, задача о перпендикуляре сводится к поиску чисел $x_1, \ldots, x_k$, которые являются решением следующей системы линейных уравнений с транспонированной матрицей Грама $G(a_1, \ldots, a_k)$ и свободными члена-

ми (f, a_i), $i = \overline{1, k}$:

$$\left\{ \sum_{j=1}^{k} x_j (a_j, a_i) = (f, a_i), \quad i = \overline{1, k}. \right. \tag{128}$$

Эта система всегда имеет решение, т. к. она эквивалентна задаче о перпендикуляре.

Отметим возможные при этом варианты.

1) Если $a_1, \ldots, a_k$ — ортонормированный базис U, то матрица Грама — единичная и система (128) имеет единственное решение $x_i = (f, a_i)$, $i = \overline{1, k}$, при этом $g = \sum\limits_{i=1}^{k} x_i a_i = \sum\limits_{i=1}^{k} (f, a_i) a_i$, $h = f - g$.

2) Если $a_1, \ldots, a_k$ — линейно независимы, то определитель Грама $\det(G(a_1, \ldots, a_k)) \neq 0$ и система (128) имеет единственное решение. Решив систему, также найдём $g = \sum\limits_{i=1}^{k} x_i a_i$, $h = f - g$.

3) Если $a_1, \ldots, a_k$ — линейно зависимы, то определитель Грама $\det(G(a_1, \ldots, a_k)) = 0$ и система (128) имеет бесконечное множество решений $x_1, \ldots, x_k$, которые соответствуют бесконечному числу разложений одного вектора g по линейно зависимой системе векторов $a_1, \ldots, a_k$: $g = \sum\limits_{i=1}^{k} x_i a_i$. А далее находим $h = f - g$. ■

7. Билинейные и квадратичные формы

7.1. Билинейные формы, их свойства и матрицы

В этом разделе рассматриваем билинейные формы в линейном пространстве E над полем P.

Определение 1. *Отображение* $B\colon E \times E \to P$ *(от двух переменных) называется* *билинейной формой*, *если*

A1) $B(f_1 + f_2, g) = B(f_1, g) + B(f_2, g)$,

A2) $B(\gamma f, g) = \gamma B(f, g)$,

A3) $B(f, g_1 + g_2,) = B(f, g_1) + B(f, g_2)$,

A4) $B(f, \delta g) = \delta B(f, g)$,

(здесь f, f_1, f_2, g, g_1, g_2 — любые векторы пространства E, γ, δ — любые числа из P).

Билинейная форма называется симметричной, если $B(f,\, g) = B(g,\, f)$ при всех $f, g \in E$.

Две формы $B_1(f,\, g)$ и $B_2(f,\, g)$ равны, если $B_1(f,\, g) = B_2(f,\, g)$ при всех $f, g \in E$.

Пример 1. Скалярное произведение (f, g) в евклидовом пространстве является симметричной билинейной формой.

Следующий пример даётся в лемме.

Лемма 7.1.1. *В n-мерном пространстве E с базисом $e = (e_1,\, e_2 \ldots, e_n)$ определим отображение $B\colon E \times E \to P$ по правилу*

$$B(f, g) = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} \alpha_i \beta_j, \tag{129}$$

для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{i=1}^{n} \beta_i e_i$, где d_{ij}, $i, j = \overline{1, n}$ — элементы произвольной матрицы $\mathcal{D}$ из $P^{n \times n}$. Тогда $B(f, g)$ является билинейной формой, а формулу (129) можно записать в матричном виде:

$$B(f, g) = [f]_e^{\top}\, \mathcal{D}\, [g]_e, \tag{130}$$

где $[f]_e^{\top} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}^{\top} = (\alpha_1\, \alpha_2, \ldots\, \alpha_n)\,, \quad [g]_e = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}.$

Доказательство. 1) $g = \sum\limits_{i=1}^{n} \beta_i e_i$, $f_1 = \sum\limits_{i=1}^{n} \alpha_i' e_i$, $f_2 = \sum\limits_{i=1}^{n} \alpha_i'' e_i \Rightarrow$ $f_1 + f_2 = \sum\limits_{i=1}^{n} (\alpha_i' + \alpha_i'') e_i$, $B(f_1,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} d_{ij} \alpha_i' \beta_j$, $B(f_2,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} d_{ij} \alpha_i'' \beta_j$, $B(f_1 + f_2,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} d_{ij} (\alpha_i' + \alpha_i'') \beta_j \Rightarrow B(f_1 + f_2,\, g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} (d_{ij} \alpha_i' \beta_j + d_{ij} \alpha_i'' \beta_j) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} d_{ij} \alpha_i' \beta_j + \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} d_{ij} \alpha_i'' \beta_j = B(f_1,\, g) + B(f_2,\, g)$.

2) $f = \sum\limits_{i=1}^{n} \alpha_i e_i,\ \ g = \sum\limits_{i=1}^{n} \beta_i e_i \Rightarrow \gamma f = \sum\limits_{i=1}^{n} (\gamma\alpha_i) e_i,$

$B(f,\, g) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} d_{ij}\alpha_i\beta_j,\ B(\gamma f,\, g) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} d_{ij}(\gamma\alpha_i)\beta_j \Rightarrow$

$\Rightarrow B(\gamma f,\, g) = \gamma \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} d_{ij}\alpha_i\beta_j = \gamma B(f,\, g).$

Аналогично проверяются свойства 3) и 4).

Совпадение (130) с (129) проверяется умножением:

$$[f]_e^{\top}\,\mathcal{D} = (\alpha_1\,\alpha_2 \ldots \alpha_n)\begin{pmatrix} d_{11} & \ldots & d_{1n} \\ d_{21} & \ldots & d_{2n} \\ \ldots & \ldots & \ldots \\ d_{n1} & \ldots & d_{nn} \end{pmatrix} = \left(\sum_{i=1}^{n}\alpha_i d_{i1}\ \ \ldots\ \ \sum_{i=1}^{n}\alpha_i d_{in}\right),$$

$$\left([f]_e^{\top}\,\mathcal{D}\right)[g]_e = \left(\sum_{i=1}^{n}\alpha_i d_{i1}\ \ \ldots\ \ \sum_{i=1}^{n}\alpha_i d_{in}\right)\begin{pmatrix} \beta_1 \\ \ldots \\ \beta_n \end{pmatrix} = \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\alpha_i d_{ij}\right)\beta_j =$$

$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\alpha_i d_{ij}\beta_j\right) = \sum_{j=1}^{n}\sum_{i=1}^{n}\alpha_i d_{ij}\beta_j = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i d_{ij}\beta_j = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}\alpha_i\beta_j. \ \blacksquare$$

Теорема 7.1.1. (Простейшие свойства билинейной формы)

1. $\forall\, m, r \in \mathbb{N},\ f_1,\ \ldots,\ f_m,\ g_1,\ \ldots,\ g_r, \in E,\ \alpha_1,\ \ldots,\ \alpha_m,\ \beta_1,\ \ldots,\ \beta_r \in P$
$$B\Big(\sum_{i=1}^{m}\alpha_i f_i,\ \sum_{j=1}^{r}\beta_j g_j\Big) = \sum_{i=1}^{m}\sum_{j=1}^{r}\alpha_i\beta_j B(f_i, g_j). \qquad (131)$$
2. $B(\theta,\, g) = 0,\ B(f,\, \theta) = 0$ *для всех* $f,\, g \in E$, *в частности,* $B(\theta,\, \theta) = 0$.

Доказательство. 1) Проведите самостоятельно с помощью математической индукции.

2) Из аксиом A2 и A4 при $\gamma = 0,\ \delta = 0$ следует $B(0f,\, g) = 0B(f,\, g)$, $B(f,\, 0g) = 0B(f,\, g) \Rightarrow B(\theta,\, g) = 0,\ B(f,\, \theta) = 0.$ $\blacksquare$

Определение 2. *Пусть* $B(f, g)$ *— билинейная форма в линейном пространстве* E *с базисом* $e = (e_1, \ldots, e_n)$.

Матрица $[B]_e \in P^{n\times n}$ *с элементами* $b_{ij} = B(e_i, e_j),\ i, j = \overline{1, n}$ *называется матрицей билинейной формы* $B(f, g)$ *в базисе* e.$\blacksquare$

Докажем существование и единственность билинейной формы, принимающей заданные значения на парах базисных векторов.

Теорема 7.1.2. *Пусть E линейное пространство над полем P и $e = (e_1, \ldots, e_n)$ базис E.*

Для любых чисел b_{ij}, $i, j = \overline{1, n}$ из P существует, и притом единственная, билинейная форма $B(f, g)$ в пространстве E, для которой

$$B(e_i, e_j) = b_{ij},\ i, j = \overline{1, n}. \tag{132}$$

Она определена формулой

$$B(f, g) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \alpha_i \beta_j, \tag{133}$$

для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{i=1}^{n} \beta_i e_i$.

Доказательство. В силу леммы 7.1.1 отображение $B(f, g)$, определённое формулой (133) является билинейной формой. Для e_i i-я координата $\alpha_i = 1$, для e_j j-я координата $\beta_j = 1$, остальные координаты равны нулю. Тогда из формулы (133) следует, что $B(e_i, e_j) = b_{ij}$, $i, j = \overline{1, n}$. Если другая билинейная форма $B_1(f, g)$ удовлетворяет условиям $B_1(e_i, e_j) = b_{ij}$, $i, j = \overline{1, n}$, то из (131) следует, что для любых $f = \sum\limits_{i=1}^{n} \alpha_i e_i$, $g = \sum\limits_{j=1}^{n} \beta_j e_j$ имеем $B_1(f, g) = B_1\big(\sum\limits_{i=1}^{n} \alpha_i e_i, \sum\limits_{j=1}^{n} \beta_j e_j\big) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \beta_j B_1(e_i, e_j) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j$, что совпадает с $B(f, g)$ в формуле (133). Следовательно, формы B и B_1 равны.

Следствие. Произвольная матрица $(b_{ij}) \in P^{n \times n}$ является матрицей единственной билинейной формы $B(f, g)$ в заданном базисе $e = (e_1, \ldots, e_n)$ пространства.

Таким образом, в лемме 7.1.1 приведён самый общий вид билинейной формы в n-мерном пространстве. ■

Теорема 7.1.3. *Матрицы билинейной формы $B(f, g)$ в базисах e и $h = eS$ связаны равенством $[B]_h = S^\top [B]_e S$.*

Доказательство. Используя (130), запишем форму в виде:

$B(f,g) = [f]_h^\top [\mathcal{B}]_h [g]_h$ и $B(f,g) = [f]_e^\top [\mathcal{B}]_e [g]_e$. С помощью замены координат $[f]_e = S[f]_h$ и $[g]_e = S[g]_h$ преобразуем второе равенство $B(f,g) = [f]_e^\top [\mathcal{B}]_e [g]_e = (S[f]_h)^\top [\mathcal{B}]_e S[g]_h = [f]_h^\top (S^\top [\mathcal{B}]_e S)[g]_h$. Далее, сравнив это равенство с равенством $B(f,g) = [f]_h^\top [\mathcal{B}]_h [g]_h$, получим $[f]_h^\top [\mathcal{B}]_h [g]_h = = [f]_h^\top (S^\top [\mathcal{B}]_e S)[g]_h$. Откуда, в силу произвольности $[f]_h^\top$, $[g]_h$, следует равенство $[B]_h = S^\top [B]_e S$. ■

Теорема 7.1.4. *Билинейная форма $B(f,g)$ в n-мерном линейном пространстве симметрична тогда и только тогда, когда её матрица симметрична в любом базисе.*

Доказательство.

Пусть $B(f,g)$ — билинейная форма в n-мерном пространстве E и $[B]_e = (b_{ij})$ — матрица этой билинейной формы в базисе $e = (e_1, \ldots, e_n)$.

Если билинейная форма симметрична, то при всех векторах f, g выполнено равенство $B(f,g) = B(g,f)$, в частности, при всех $i, j = \overline{1,n}$ $B(e_i, e_j) = B(e_j, e_i)$, т.е. $b_{ij} = b_{ji}$. Значит, матрица симметрична.

Обратно, пусть в базисе e матрица билинейной формы симметрична, т.е. $b_{ij} = b_{ji}$ при всех $i, j = \overline{1,n}$. Для произвольных $f = \sum\limits_{i=1}^{n} \alpha_i e_i$ $g = \sum\limits_{i=1}^{n} \beta_i e_i$ из формулы (133) получаем $B(f,g) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j$ и $B(g,f) = = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \beta_i \alpha_j$. Далее из $b_{ij} = b_{ji}$ имеем $B(g,f) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ji} \alpha_j \beta_i$. Так как $\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ij} \alpha_i \beta_j = \sum\limits_{j=1}^{n} \sum\limits_{i=1}^{n} b_{ji} \alpha_j \beta_i = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} b_{ji} \alpha_j \beta_i$, то $B(f,g) = B(g,f)$. ■

7.2. Квадратичные формы, их свойства и матрицы. Канонический вид квадратичной формы. Приведение квадратичной формы к каноническому виду

Определение 1. *Пусть $B(f,g)$ — симметричная билинейная форма в пространстве E над полем P.*

Квадратичной формой называется отображение $K\colon E \to P$, определённое для произвольного $f \in E$ формулой $K(f) = B(f,f)$.

При этом, форма $B(f,g)$ называется полярной билинейной формой к квадратичной форме $K(f)$.

Матрицей квадратичной формы $K(f)$ в базисе $e=(e_1,\ldots,e_n)$ называется матрица полярной к ней билинейной формы $B(f,g)$ в том же базисе. Обозначим эту матрицу $[K]_e$. Таким образом, $[K]_e=[B]_e$. ■

Запишем в координатном виде квадратичную форму.

Лемма 7.2.1. *Пусть $K(f)$ — квадратичная форма с матрицей $[K]_e=(b_{ij})$ в базисе $e=(e_1,\ldots,e_n)$ и произвольный вектор f задан своими координатами $f=\sum\limits_{i=1}^{n} x_ie_i$.*

Тогда

$$K(f)=\sum_{i,j=1}^{n} b_{ij}x_ix_j=\sum_{i=1}^{n} b_{ii}x_i^2+2\sum_{i<j} b_{ij}x_ix_j=b_{11}x_1^2+2b_{12}x_1x_2+\ldots+$$
$$+2b_{1n}x_1x_n++b_{22}x_2^2+2b_{23}x_2x_3+\ldots+2b_{2n}x_2x_n+\ldots+2b_{n-1,n}x_{n-1}x_n+$$
$$+b_{nn}x_n^2. \qquad (134)$$

Доказательство. Следует из (129) и симметричности матрицы. ■

Меняя базис пространства, можно изменить вид матрицы квадратичной формы и упростить её.

Определение 2. *Базис $e=(e_1,\ldots,e_n)$ называется каноническим базисом квадратичной формы, если квадратичная форма K имеет в этом базисе диагональную матрицу $[K]_e=diag(d_1,d_2,\ldots,d_n)$.*

В каноническом базисе квадратичная форма согласно (134) имеет вид $K(f)=d_1x_1^2+d_2x_2^2+\ldots+d_nx_n^2$, который называется каноническим видом, а числа $d_1,d_2,\ldots,d_n$ — каноническими коэффициентами. Канонический вид называют также суммой квадратов. ■

Количестве ненулевых канонических коэффициентов совпадает с рангом матрицы квадратичной формы.

Теорема 7.2.1 (О приведении к каноническому виду).

Всякую квадратичную форму невырожденным преобразованием координат можно привести к каноническому виду.

Доказательство. Пусть дана квадратичная форма $K = \sum\limits_{i,j=1}^{n} b_{ij}x_ix_j$ от n координат $x_1, x_2, \ldots, x_n$. Постараемся представить K в виде суммы двух слагаемых, первое – это квадрат выражения, содержащего все слагаемые с одной их координат, второе — квадратичная формы от остальных $n-1$ координат. Эта цель легко достигается в случае, когда среди диагональных коэффициентов матрицы $b_{11}, b_{22}, \ldots, b_{nn}$, стоящих при квадратах координат, есть отличные от нуля.

Пусть, например, $b_{11} \neq 0$. Из (134) выделим и преобразуем все слагаемые, содержащие x_1 :

$$K = \sum_{i=1}^{n} b_{ii}x_i^2 + 2\sum_{i<j} b_{ij}x_ix_j =$$

$$= \big(b_{11}x_1^2 + 2\sum_{1<j} b_{1j}x_1x_j\big) + \big(\sum_{i=2}^{n} b_{ii}x_i^2 + 2\sum_{1<i<j} b_{ij}x_ix_j\big) = V + W.$$

$$V = b_{11}x_1^2 + 2b_{12}x_1x_2 + \ldots + 2b_{1n}x_1x_n = b_{11}x_1^2 + 2x_1(b_{12}x_2 + \ldots + b_{1n}x_n) =$$

$$= \frac{1}{b_{11}}\Big((b_{11}x_1)^2 + 2(b_{11}x_1)(b_{12}x_2 + \ldots + b_{1n}x_n)\Big) = \frac{1}{b_{11}}\Big((b_{11}x_1)^2 +$$

$$+2(b_{11}x_1)(b_{12}x_2+\ldots+b_{1n}x_n)+(b_{12}x_2+\ldots+b_{1n}x_n)^2-(b_{12}x_2+\ldots+b_{1n}x_n)^2\Big) =$$

$$= \frac{1}{b_{11}}\Big((b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n)^2 - (b_{12}x_2 + \ldots + b_{1n}x_n)^2\Big) =$$

$$= \frac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n)^2 - \frac{1}{b_{11}}(b_{12}x_2 + \ldots + b_{1n}x_n)^2.$$

Итак, $K = \dfrac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n)^2 + \Big(-\dfrac{1}{b_{11}}(b_{12}x_2 + \ldots + b_{1n}x_n)^2 +$ $+\sum\limits_{i=2}^{n} b_{ii}x_i^2 + 2\sum\limits_{1<i<j} b_{ij}x_ix_j\Big) = \dfrac{1}{b_{11}}(b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n)^2 + K_1$, где K_1 является квадратичной формой от координат $x_2, \ldots, x_n$.

Если мы теперь введём обозначения

$$\begin{aligned} &y_1 = b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n, \\ &y_i = x_i, \ i = \overline{2, n}, \end{aligned} \tag{135}$$

то получим

$$K = \frac{1}{b_{11}}y_1^2 + K_1,$$

где K_1 будет теперь квадратичной формой от координат $y_2, \ldots, y_n$. Равенства (135) определяют линейное преобразование координат

$$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

с невырожденной матрицей, т.к. её определитель равен $b_{11} \neq 0$. В формулах преобразования координат старые координаты выражаются через новые координаты

$$x_1 = \frac{1}{b_{11}} y_1 - \frac{b_{12}}{b_{11}} y_2 - \dots - \frac{b_{1n}}{b_{11}} y_n, \quad x_i = y_i, \ i = \overline{2, n}.$$

Эти равенства определяют линейное преобразование координат с обратной матрицей, которая также невырождена.

Теперь рассмотрим случай, когда все диагональные элементы матрицы квадратичной формы равны нулю $a_{11} = a_{22} = \dots = a_{nn} = 0$. Если все элементы матрицы равны нулю, то она имеет канонический вид с нулевыми каноническими коэффициентами. Поэтому предположим, что среди элементов матрицы есть ненулевые, например, $b_{12} \neq 0$. Тогда $K = 2b_{12}x_1x_2 + K_2$, где каждое слагаемое в K_2 содержит хотя бы одну координату $x_3, \dots, x_n$. Совершим линейное преобразование координат

$$\begin{aligned} x_1 &= y_1 - y_2, \\ x_2 &= y_1 + y_2, \\ x_i &= y_i, i = \overline{3, n}, \end{aligned} \tag{137}$$

Оно будет невырожденным, т. к. определитель

$$\begin{vmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & \dots\dots\dots & & \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} 1 & \dots & 0 \\ & \dots\dots & \\ 0 & \dots & 1 \end{vmatrix} = 2 \neq 0.$$

В результате член $2b_{12}x_1x_2$ примет вид

$$2b_{12}x_1x_2 = 2b_{12}(y_1 - y_2)(y_1 + y_2) = 2b_{12}y_1^2 - 2b_{12}y_2^2,$$

т.е. в форме появятся два квадрата $b_{12}y_1^2$ и $-2b_{12}y_2^2$ с ненулевыми коэффициентами, причём они не могут сократиться ни с одним из остальных членов, т.к. в каждый из них входит хотя бы одна координата $y_3, \dots, y_n$. Теперь мы находимся в условиях первого случая и можем ещё одним

невырожденным преобразованием привести её к виду (136).

Применяя это рассуждение к квадратичной форме с меньшим числом координат, мы за конечное число шагов получим канонический вид. Выполнив при этом все преобразования координат, мы получим формулы, связывающие начальные и конечные координаты . ■

Задачи.

Задача 7.2.1. *Найдите канонический вид и невырожденное линейное преобразование, приводящее к этому виду, следующей квадратичной формы* $K = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 12x_2x_3$.

Р е ш е н и е. 1-й шаг. Отметим, что квадратичная форма K не является суммой квадратов и в ней присутствует ненулевой квадрат, например, $2x_1^2$. Поэтому выделим из K квадрат выражения, содержащего все слагаемые с координатой x_1 : $K = (2x_1^2+4x_1x_2-4x_1x_3)+(5x_2^2+5x_3^2-12x_2x_3)$. Обозначим $V = 2x_1^2 + 4x_1x_2 - 4x_1x_3$, $W = 5x_2^2 + 5x_3^2 - 12x_2x_3$, тогда $K = V + W$. Далее преобразуем V.

$$V = \frac{1}{2}\big(4x_1^2 + \underline{8x_1x_2 - 8x_1x_3}\big) = \frac{1}{2}\big((2x_1)^2 + 2(2x_1)(2x_2 - 2x_3)\big) =$$
$$= \frac{1}{2}\Big(\big((2x_1)^2 + 2(2x_1)(2x_2 - 2x_3) + (\underline{2x_2 - 2x_3})^2\big) - (\underline{2x_2 - 2x_3})^2\Big) =$$
$$= \frac{1}{2}\Big(\big((2x_1) + (2x_2 - 2x_3)\big)^2 - (2x_2 - 2x_3)^2\Big) =$$
$$= \frac{1}{2}\big(2x_1 + 2x_2 - 2x_3\big)^2 - \frac{1}{2}(2x_2 - 2x_3)^2 =$$
$$= \frac{2^2}{2}(x_1 + x_2 - x_3)^2 - \frac{2^2}{2}(x_2 - x_3)^2 =$$
$$= 2(x_1 + x_2 - x_3)^2 - 2(x_2^2 - 2x_2x_3 + x_3^2) =$$
$$= 2(x_1 + x_2 - x_3)^2 - 2x_2^2 + 4x_2x_3 - 2x_3^2.$$

Отсюда $K = V + W = 2(x_1 + x_2 - x_3)^2 - 2x_2^2 + 4x_2x_3 - 2x_3^2 + 5x_2^2 + 5x_3^2 - 12x_2x_3 = 2(x_1 + x_2 - x_3)^2 + (3x_2^2 - 8x_2x_3 + 3x_3^2)$. Здесь только первое выражение $2(x_1+x_2-x_3)^2$ содержит переменную x_1. Обозначим второе выражение $K_1 = 3x_2^2-8x_2x_3+3x_3^2$, оно является квадратичной формой от переменных x_2, x_3. После преобразования квадратичная форма приняла вид $K = 2(x_1 + x_2 - x_3)^2 + K_1$.

2-й шаг. Отметим, что квадратичная форма K_1 не является суммой квадратов и в ней присутствует ненулевой квадрат, например, $3x_2^2$. Поэтому выделим из K_1 квадрат выражения, содержащего все слагаемые с координатой x_2 : $K_1 = (3x_2^2 - 8x_2x_3) + 3x_3^2$. Обозначим $V_1 = 3x_2^2 - 8x_2x_3$, $W_1 = 3x_3^2$, тогда $K_1 = V_1 + W_1$. Далее преобразуем V_1.
$V_1 = \frac{1}{3}(9x_2^2 - 24x_2x_3) = \frac{1}{3}\Big(\big((3x_2)^2 - 2(3x_2)(4x_3) + (4x_3)^2\big) - (4x_3)^2\Big) =$
$= \frac{1}{3}(3x_2 - 4x_3)^2 - \frac{16}{3}x_3^2$. Отсюда $K_1 = V_1 + W_1 = \frac{1}{3}(3x_2 - 4x_3)^2 - \frac{16}{3}x_3^2 +$
$+ 3x_3^2 = \frac{1}{3}(3x_2 - 4x_3)^2 - \frac{7}{3}x_3^2$. После преобразования квадратичная форма K_1 является алгебраической суммой квадратов, поэтому выделение квадратов закончено. Запишем преобразованный вид квадратичной формы $K = 2(x_1 + x_2 - x_3)^2 + K_1 = 2(x_1 + x_2 - x_3)^2 + \frac{1}{3}(3x_2 - 4x_3)^2 - \frac{7}{3}x_3^2$.
Введём обозначения

$$\begin{cases} y_1 = x_1 + x_2 - x_3, \\ y_2 = 3x_2 - 4x_3, \\ y_3 = x_3. \end{cases} \tag{138}$$

Это будет невырожденное преобразование координат, т.к. определитель

$$\begin{vmatrix} 1 & 1 & -1 \\ 0 & 3 & -4 \\ 0 & 0 & 1 \end{vmatrix} = 3 \neq 0.$$

Тогда квадратичную форму можно записать в каноническом виде в новых координатах y_1, y_2, y_3 :

$$K = 2y_1^2 - \frac{1}{3}y_2^2 - \frac{7}{3}y_3^2.$$

Невырожденное линейное преобразование, приводящее к этому виду, получим, выразив из (138) старые координаты x_1, x_2, x_3 через новые:

$$\begin{cases} x_3 = y_3, \\ 3x_2 = y_2 + 4x_3, \\ x_1 = y_1 - x_2 + x_3, \end{cases} \Rightarrow \begin{cases} x_3 = y_3, \\ 3x_2 = y_2 + 4y_3, \\ x_1 = y_1 - x_2 + y_3, \end{cases} \Rightarrow \begin{cases} x_3 = y_3, \\ x_2 = \dfrac{1}{3}y_2 + \dfrac{4}{3}y_3, \\ x_1 = y_1 - x_2 + y_3, \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x_3 = y_3, \\ x_2 = \dfrac{1}{3}y_2 + \dfrac{4}{3}y_3, \\ x_1 = y_1 - \dfrac{1}{3}y_2 - \dfrac{4}{3}y_3 + y_3, \end{cases} \Rightarrow \begin{cases} x_3 = y_3, \\ x_2 = \dfrac{1}{3}y_2 + \dfrac{4}{3}y_3, \\ x_1 = y_1 - \dfrac{1}{3}y_2 - \dfrac{1}{3}y_3, \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x_1 = y_1 - \dfrac{1}{3}y_2 - \dfrac{1}{3}y_3, \\ x_2 = \dfrac{1}{3}y_2 + \dfrac{4}{3}y_3, \\ x_3 = y_3. \end{cases}$$